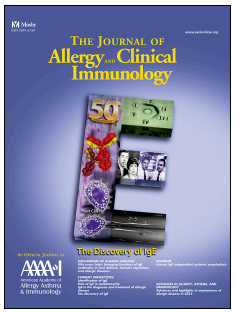
the 23 and Me Research Team, & the collaborators of the SHARE study (2018). Eleven loci with new reproducible genetic associations with allergic disease risk. *Journal of Allergy and Clinical Immunology*. https://doi.org/10.1016/j.jaci.2018.03.012

Accepted Manuscript



Eleven loci with new reproducible genetic associations with allergic disease risk

Manuel AR. Ferreira, PhD, Judith M. Vonk, PhD, Hansjörg Baurecht, PhD, Ingo Marenholz, PhD, Chao Tian, PhD, Joshua D. Hoffman, PhD, Quinta Helmer, PhD, Annika Tillander, PhD, Vilhelmina Ullemar, PhD, Yi Lu, PhD, Franz Rüschendorf, PhD, David A. Hinds, PhD, Norbert Hübner, MD, Stephan Weidinger, MD, Patrik KE. Magnusson, PhD, Eric Jorgenson, PhD, Young-Ae Lee, MD, Dorret I. Boomsma, PhD, Robert Karlsson, PhD, Catarina Almqvist, MD, Gerard H. Koppelman, MD, Lavinia Paternoster, PhD

PII: S0091-6749(18)30558-X

DOI: 10.1016/j.jaci.2018.03.012

Reference: YMAI 13380

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 24 November 2017

Revised Date: 1 February 2018

Accepted Date: 19 March 2018

Please cite this article as: Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, Helmer Q, Tillander A, Ullemar V, Lu Y, Rüschendorf F, the 23andMe Research Team, collaborators of the SHARE study, Hinds DA, Hübner N, Weidinger S, Magnusson PK, Jorgenson E, Lee Y-A, Boomsma DI, Karlsson R, Almqvist C, Koppelman GH, Paternoster L, Eleven loci with new reproducible genetic associations with allergic disease risk, *Journal of Allergy and Clinical Immunology* (2018), doi: 10.1016/j.jaci.2018.03.012.

Eleven loci with new reproducible genetic associations with allergic disease risk

Manuel AR Ferreira, PhD [1], Judith M Vonk, PhD [2], Hansjörg Baurecht, PhD [3], Ingo Marenholz, PhD [4,5], Chao Tian, PhD [6], Joshua D Hoffman, PhD [7], Quinta Helmer, PhD [8], Annika Tillander, PhD [9], Vilhelmina Ullemar, PhD [9], Yi Lu, PhD [9], Franz Rüschendorf, PhD [4], the 23andMe Research Team[6], collaborators of the SHARE study[10], David A Hinds, PhD [6], Norbert Hübner, MD [4], Stephan Weidinger, MD[3], Patrik KE Magnusson, PhD [9], Eric Jorgenson, PhD [11], Young-Ae Lee, MD [4,5], Dorret I Boomsma, PhD [8], Robert Karlsson, PhD [9], Catarina Almqvist, MD [9,12], Gerard H Koppelman, MD [13] and Lavinia Paternoster, PhD [14]

Affiliations

[1] Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Australia

[2] Epidemiology, University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma and COPD, Groningen, the Netherlands

[3] Department of Dermatology, Allergology and Venereology, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany

[4] Max Delbrück Center (MDC) for Molecular Medicine, Berlin, Germany

[5] Clinic for Pediatric Allergy, Experimental and Clinical Research Center of Charité Universitätsmedizin Berlin and Max Delbrück Center, Berlin, Germany

[6] 23andMe, Inc., Mountain View, California, USA

[7] Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA

[8] Department Biological Psychology, Netherlands Twin Register , Vrije University, Amsterdam, The

Netherlands

[9] Department of Medical Epidemiology and Biostatistics and the Swedish Twin Registry, Karolinska Institutet, Stockholm, Sweden

[10] Collaborators of the SHARE study are listed in this article's **Online Repository**

[11] Division of Research, Kaiser Permanente Northern California, Oakland, California, USA

[12] Pediatric Allergy and Pulmonology Unit at Astrid Lindgren Children's Hospital, Karolinska University Hospital, Stockholm, Sweden

[13] University of Groningen, University Medical Center Groningen, Beatrix Children's Hospital, Pediatric Pulmonology and Pediatric Allergology, and University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma and COPD, Groningen, the Netherlands

[14] MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK

Corresponding author:

Manuel A R Ferreira, PhD

QIMR Berghofer Medical Research Institute

Locked Bag 2000, Royal Brisbane Hospital,

Herston QLD 4029, Australia

Phone: +61 7 3845 3552

Fax: +61 7 3362 0101

Email: manuel.ferreira@qimrberghofer.edu.au

ABSTRACT

Background: A recent genome-wide association study (GWAS) identified 99 loci that contain genetic risk variants shared between asthma, hay fever and eczema. Many more risk loci shared between these common allergic diseases remain to be discovered, which could point to new therapeutic opportunities.

Objective: To identify novel risk loci shared between asthma, hay fever and eczema by applying a gene-based test of association to results from a published GWAS that included data from 360,838 individuals.

Methods: We used approximate conditional analysis to adjust the results from the published GWAS for the effects of the top risk variants identified in that study. We then analysed the adjusted GWAS results with the EUGENE gene-based approach, which combines evidence for association with disease risk across regulatory variants identified in different tissues. Novel gene-based associations were followed up in an independent sample of 233,898 individuals from the UK Biobank study.

Results: Of the 19,432 genes tested, 30 had a significant gene-based association at a Bonferroni-corrected *P*-value of 2.5×10^{-6}. Of these, 20 were also significantly associated ($P<0.05/30=0.0016$) with disease risk in the replication sample, including 19 that were located in 11 loci not reported to contain allergy risk variants in previous GWAS. Amongst these were nine genes with a known function that is directly relevant to allergic disease: *FOSL2*, *VPRBP*, *IPCEF1*, *PRR5L*, *NCF4*, *APOBR*, *IL27*, *ATXN2L* and *LAT*. For four genes (e.g. *ATXN2L*), a genetically-determined decrease in gene expression was associated with decreased allergy risk, and therefore drugs that inhibit gene expression or function are predicted to ameliorate disease symptoms. The opposite directional effect was observed for 14 genes, including *IL27*, a cytokine known to suppress Th2 responses.

Conclusion: Using a gene-based approach, we identified 11 risk loci for allergic disease that were not reported in previous GWAS. Functional studies that investigate the contribution of the 19 associated genes to the pathophysiology of allergic disease and assess their therapeutic potential are warranted.

INTRODUCTION

The strong genetic correlations between asthma, hay fever and eczema estimated from twin studies[1-5], in combination with the highly polygenic architecture of these diseases, predict that many hundreds if not thousands of genetic risk factors are shared between these three common allergic diseases. Motivated by this prediction, we recently performed a genome-wide association study (GWAS) designed to identify genetic risk variants that are shared between asthma, hay fever and eczema [6]. In that GWAS, cases (n=180,129) were defined as individuals who reported having suffered from one or more allergic disease, while controls (n=180,709) were individuals who reported never having suffered from any of these diseases. We identified 136 single nucleotide polymorphisms (SNPs) located in 99 loci (*i.e.* genomic regions located > 1Mb apart) that were independently associated with disease risk at a genome-wide significance-threshold of $P<3x10^{-8}$, a threshold that corrects for the number of SNPs tested [7]. In our study, we often observed that multiple independent genetic variants within one locus contributed to disease risk – this was the case for 18 of the 99 risk loci identified.

Larger GWAS of this multi-disease phenotype are underway and are expected to identify more risk variants shared between asthma, hay fever and eczema. Here, we report results from another approach that increases power to identify novel risk loci: gene-based instead of SNP-based association analysis. Different gene-based tests have been developed, including VEGAS [8], MAGMA [9] and fastBAT [10]. For each gene in the genome, these methods combine in a single test the evidence for association with a disease across multiple SNPs, which are typically selected because they are located in or near that gene (*e.g.* within 100 kb). These methods improve power over the alternative approach of testing each SNP at a time (as is done in a GWAS) when multiple SNPs near a gene are independently associated with disease risk. However, not all SNPs near a gene are directly relevant to its function. For example, only some SNPs influence variation in gene expression levels – these are commonly referred to as expression quantitative trait loci (eQTL). On the other hand, eQTL have a greater probability of

being associated with common diseases and traits [11]. Motivated by these observations, an additional suite of gene-based methods has been developed recently that only include in the association test functional SNPs, such as eQTL. These include, for example, PrediXcan [12], EUGENE [13] and S-PrediXcan [14]. We developed EUGENE because it was not possible with other methods to combine in the same association test information from eQTL identified in different tissues. This feature is important because multiple tissue types play a role in allergic disease pathophysiology and tissue-specific eQTL are common [15]. Furthermore, EUGENE also considers regulatory variants with different mechanisms of action, for example, variants that affect splicing but not overall transcription levels. This increased resolution is expected to increase our ability to identify genes that are causally-related to common diseases [16].

The aims of the present study were to (1) identify novel risk loci shared between asthma, hay fever and eczema by applying EUGENE to results from our previous GWAS [6]; and (2) follow up the top gene-based associations in an independent replication sample ascertained from the UK Biobank study[17].

METHODS

Adjusting GWAS results for the effects of genome-wide significant SNPs

The starting point for this study was a GWAS of allergic disease reported recently by Ferreira et al. [6], which included 360,838 individuals from 13 studies: UK Biobank, 23andMe, GERA, CATSS, NTR, LifeLines, TWINGENE, ALSPAC, SALTY, GENEVA, AAGC, GENUFAD-SHIP-1 and GENUFAD-SHIP-2 (demographics in **Table E1** in this article's **Online Repository**). In that study, single-SNP results were corrected for an inflation factor that reflected technical biases and/or population stratification (specifically, an LD-score regression intercept[18] of 1.04). After that correction, there were 136 variants independently associated with allergic disease in Ferreira et al. [6]. Because our aim was to identify new allergy risk loci, we first applied approximate conditional analyses as implemented in the GCTA tool [19] to the summary statistics of Ferreira et al. [6] to adjust the single-SNP results for the effects of the 136 risk variants (a detailed justification for performing conditional analysis prior to the gene-based analysis is provided in the Online Repository). In the conditional analysis, linkage disequilibrium (LD) between SNPs was estimated using genotype data from 5,000 individuals from the UK Biobank study[17].

Gene-based analysis of the adjusted GWAS results

To identify novel risk loci shared between asthma, hay fever and eczema, we analysed the adjusted GWAS results with EUGENE[13], a gene-based approach that is applicable to summary statistics (*i.e.* it does not require individual-level genetic data for subjects included in the GWAS) and combines evidence for association with disease risk across eQTL identified in different tissues. The latter feature is important because multiple tissue types play a role in allergic disease pathophysiology and tissue-specific eQTL are common [15].

We identified eQTL based on information from 39 published eQTL studies conducted in 19 tissues or cell types relevant to allergic disease (**Table E2** in this article's **Online Repository**). For each eQTL study, we (i) downloaded the original publication tables containing results for the eQTL reported; (ii) extracted the SNP identifier, gene name, association P-value and directional effect (if available; beta/z-score and effect allele) for all reported eQTL; and (iii) excluded eQTL located >1 Mb of the respective gene (i.e. *trans* eQTL) and/or with an association $P>2.3x10^{-9}$, a conservative threshold that corrects for 21,742 genes in the genome, each tested for association with 1,000 SNPs (as suggested by others [20-22]). We did not include *trans* eQTLs in the analysis because often these are thought to involve indirect effects [23], for example, where a SNP influences the expression of a gene in *cis*, which in turn affects the expression of many other genes in *trans*.

Having identified a list of *cis* eQTL for a given gene from published studies, we then reduced that list to a set of eQTL in low LD with each other (specifically, with an $r^2<0.1$) using the --clump procedure in PLINK v1.90 [24]. For a given gene, we refer to these as "independent eQTL", although we recognise that some pairs of eQTL will not be in linkage equilibrium. LD was estimated based on genetic data from individuals of European descent from the 1000 Genomes Project 23 (n=294, release 20130502_v5a). Clumping was not performed separately for each tissue or study, but rather applied to the overall list of eQTL obtained after considering information from all tissues/studies. If an eQTL was identified in multiple tissues/studies, the clumping procedure was performed using the smallest *P*-value reported for that eQTL across all tissues/studies. A file (ASTHMA.20170517.eqtl.proxies.list) containing the independent eQTL identified per gene is available at https://genepi.qimr.edu.au/staff/manuelF/eugene/main.html.

For each gene, EUGENE extracts single-SNP association results for each independent eQTL (or, if not available, for a proxy with $r^2>0.8$) from the GWAS summary statistics, sums the association chi-square values across those eQTL, and estimates the significance of the resulting sum test statistic

using Satterthwaite's approximation, which accounts for the LD between eQTL. This approximation was originally implemented by Bakshi et al. [10] in the GCTA-fastBAT module and is now also available in EUGENE. LD between eQTL was estimated based on data from 294 Europeans from the 1000 Genomes Project (release 20130502_v5a). The significance threshold required to achieve experiment-wide significance was set at $P < 0.05 / N$ genes tested.

Replication of significant gene-based associations in an independent sample

To confirm novel gene-based associations, we analysed an independent sample of unrelated individuals of European descent from the UK Biobank study [17]. The approach used to select individuals for analysis was very similar to that described in detail previously [6]. Briefly, we (1) downloaded array (805,426 variants) and imputed (92,693,895 variants) genetic data for the entire UK Biobank study, comprising 488,377 individuals, in June 2017; (2) pruned the array data down to a set of 29,446 independent SNPs that had comparable (P>0.005) allele frequencies between Europeans of the 1000 Genomes Project (CEU, FIN, GBR and TSI groups) and the UK Biobank individuals of European descent included in the Ferreira et al. GWAS [6]; (3) merged the pruned dataset with data from 1,092 individuals of known ancestry from the 1000 Genomes Project (release 20130502_v5a); (4) performed multi-dimensional scaling (MDS) analysis of identity-by-state allele sharing separately for each of 32 groups of ~16,000 individuals (to be computationally feasible), including those from the 1000 Genomes Project; (5) identified and removed individuals who did not cluster closely (within 5 standard deviations of MDS components 1 and 2) to individuals of European ancestry from the 1000 Genomes Project, resulting in 461,885 individuals; (6) identified and removed any individuals included in, or related to (*i.e.* with a kinship coefficient that indicates 3rd degree relatedness or closer; based on file ukb1007_rel_s488374.dat), the 138,354 individuals of the UK Biobank study included in Ferreira et al. [6], as well as individuals (i) with genetically-inferred sex different from self-reported sex, (ii) who were

184 outliers for SNP missingness or genome-wide heterozygosity levels, and/or (iii) with >10 third-degree
185 relatives, resulting in 244,395 individuals.

186 For each individual, allergic disease status was defined as previously described for the UK
187 Biobank study in detail [6]. To classify asthma status, we combined information from three sources: (1)
188 touchscreen questionnaire (data-field 6152); (2) Non-cancer illness code, self-reported during verbal
189 interview (data-field 20002); (3) main (data-field 41202) and secondary (data-field 41204) ICD10
190 diagnoses. Specifically, inclusion criteria for cases were: (i) a report of "Asthma" in field 6152 and a
191 code for asthma (1111) in field 20002; or (ii) an ICD10 code for asthma in fields 41202 or 41204.
192 Exclusion criteria for cases were: (i) a report of COPD in fields 6152 or 20002; or (ii) other self-
193 reported respiratory diseases in field 20002. Inclusion criteria for controls were no report of asthma in
194 fields 6152, 20002, 41202 and 41204. Exclusion criteria for controls were the same as for cases (no
195 COPD or other self-reported respiratory diseases). To classify hay fever status, we used the same three
196 sources of information. Specifically, inclusion criteria for cases were: (i) a report of "Hay fever, allergic
197 rhinitis or eczema" in field 6152 and a code for hay fever (1387) in field 20002; or (ii) an ICD10 code
198 for hay fever in fields 41202 or 41204. Inclusion criteria for controls were no report of hay fever in
199 fields 6152, 20002, 41202 and 41204. The eczema phenotype was created very similarly to the hay
200 fever phenotype. Inclusion criteria for cases were: (i) a report of "Hay fever, allergic rhinitis or
201 eczema" in field 6152 and a code for eczema (1452) in field 20002; or (ii) an ICD10 code for eczema
202 in fields 41202 or 41204. Inclusion criteria for controls were no report of eczema in fields 6152, 20002,
203 41202 and 41204. To create the overall allergic disease phenotype used for analysis, cases were
204 individuals classified as suffering from at least one condition (asthma and/or hay fever and/or eczema),
205 as described above. Controls were individuals classified as not having suffered from all three
206 conditions. Using this approach, the 244,395 individuals selected for analysis (see above) included
207 71,807 allergic disease cases, 162,091 allergic disease controls and 10,497 individuals with a missing

phenotype.

SNPTEST was then used to test the association between disease status and imputed genotype data for eQTL of genes selected for replication; age, sex and SNP chip were included as covariates. We only analysed SNPs imputed based on the Haplotype Reference Consortium panel, given that variants imputed from the UK10K + 1000 Genomes panel were not mapped correctly. We also tested the association with 1.2 million HapMap3 SNPs in order to estimate the degree of inflation in test statistics arising because of unaccounted technical biases, using the LD Score approach[25]. We observed an LD Score intercept of 1.09, which was used to adjust the association results for the eQTL tested. Lastly, EUGENE was used as described above to perform the gene-based analysis for all genes selected for replication.

Association analyses contrasting individuals suffering from a single allergic disease

The case-control phenotype analysed in our GWAS[6] combined information from asthma, hay fever and eczema, and so was expected to improve power to identify risk variants shared between, but not specific to any of, the three diseases [26]. To understand if gene-based associations discovered through the analysis of that multiple-disease phenotype were indeed likely to represent risk factors shared across allergic diseases, we performed case-only association analyses as described in detail previously [6]. First, we tested the association between eQTL of selected genes with three phenotypes that contrasted three non-overlapping groups of adults who suffered from a single allergic disease: asthma only cases (g1; *n*=12,268) vs. hay fever only cases (g2; *n*=33,305); asthma only cases (g1) vs. eczema only cases (g3; *n*=6,276); and hay fever only cases (g2) vs. eczema only cases (g3). For a given eQTL, results from these analyses indicate if the risk allele is more (odds ratio [OR] >1) or less (OR<1) common in *e.g.* group 1 (g1) when compared to group 2 (g2). For example, if an eQTL contributed similarly to the risks of asthma and hay fever but not eczema, then one would expect an OR~1 in the asthma only vs.

hay fever only comparison, but an OR>1 in the asthma vs. eczema and hay fever vs. eczema analyses. Second, for each gene and phenotype tested, we combined the association results across eQTL using EUGENE as we did in the analysis of the adjusted GWAS results described above.

This study was approved by the Human Ethics Committee of the QIMR Berghofer Medical Research Institute.

RESULTS

Identification of novel risk loci for allergic disease through gene-based association analysis

An overview of the analytical approach used is shown in **Figure 1**. To identify novel risk loci shared between asthma, hay fever and eczema, we first adjusted the association results from Ferreira et al. [6] for the 136 genome-wide significant SNPs (*i.e.* with $P<3x10^{-8}$) identified in that study using approximate conditional analysis (**Figure 2A**). In the resulting adjusted GWAS, as expected there were no SNPs associated with disease risk at $P<3x10^{-8}$ and located >1Mb from the loci reported in Ferreira et al. [6]. On the other hand, four SNPs located in loci reported in that study (in/near *CADM3*, *SLC39A8*, *LRRC43* and *KLF5*) were genome-wide significant in the conditional but not in the original analyses, consistent with the presence of additional secondary association signals at those established risk loci (**Figure 2B**). Importantly, there was an enrichment in significant SNP associations in the adjusted GWAS (**Figure E1 in the Online Repository**), suggesting that many of these associations are likely to represent true-positive findings.

To identify loci that were likely to contain true-positive associations, we applied the EUGENE gene-based approach [13] to the adjusted GWAS results. Specifically, we tested the association between disease risk and 19,432 genes (or other types of transcripts, such as long non-coding RNAs [lncRNA]) that were reported to have one or more independent eQTL in 19 tissues or cell types relevant to allergic disease (**Table E2** in this article's **Online Repository**), including whole blood, lung, skin and individual immune cell types.

We identified 30 significant gene-based associations at a Bonferroni-corrected *P*-value of $2.5x10^{-6}$, which were located in 18 loci (**Table 1**). The specific eQTL included in the gene-based test for each of these 30 genes are listed in **Table E3** in this article's **Online Repository**. For 21 genes, the association *P*-value obtained with the gene-based test was more significant than the *P*-value obtained with the individual eQTL most associated with disease risk (**Figure 3**), indicating that multiple eQTL

for the same gene were associated with disease risk (range 2 to 7 associated eQTL per gene; **Table 1**). For eight genes, the difference in significance between the most associated individual eQTL and the gene-based test exceeded one order of magnitude (**Figure 3**). The most extreme example of this was the *SPNS1* gene on chromosome 16p11.2 (gene-based $P=3.5x10^{-9}$ versus best individual eQTL $P=4.8x10^{-6}$), for which 6 of the 15 eQTL tested (identified in five tissues) were nominally associated with disease risk (**Table E4** in this article's **Online Repository**).

Replication of significant gene-based associations in an independent sample

Next, we performed a replication study to determine which of the 30 significant gene-based associations were likely to represent true-positive findings. To this end, we first identified 71,807 cases and 162,091 controls genotyped by the UK Biobank study[17] who were unrelated to individuals from our initial GWAS [6]. We then used EUGENE to test the association in this independent sample between case-control status and each of the 30 genes identified above.

There were 20 significant gene-based associations at a conservative significance threshold that accounts for the 30 genes tested ($P<0.05/30=0.0016$; **Table 1**). These included 19 genes located in 11 loci not implicated in allergic disease in previous GWAS: *OR10J5* (chromosome 1q23.2); *RP11-534L20.5* (1q32.1); *FOSL2* (2p23.2); *RBM15B* and *VPRBP* (3p21.2); *IPCEF1* (6q25.2); *AC004893.11* (7q22.1); *PRR5L* (11p13); *NSMCE1* (16p12.1); *SPNS1* and seven other nearby genes (16p11.2); *PVALB* and *NCF4* (22q12.3). Of note, nine of these genes have a known function that is directly relevant to the pathophysiology of allergic disease (**Table 2**).

Predicted directional effect of gene expression on disease risk

Because the gene-based approach used focuses exclusively on eQTL, which in turn are associated with the expression of specific genes, we were able to identify in a single analysis both novel risk loci as

well as the likely gene(s) underlying each association. Furthermore, often (but not always) the directional effect of an eQTL on gene expression can be obtained from published eQTL studies. Based on this information, for each gene, we determined if the allergy-protective allele of eQTL included in the gene-based test was associated with increased or decreased gene expression. This is important because drugs that mimic the directional effect of the allergy-protective allele on gene expression might be expected to attenuate (rather than exacerbate) allergic disease symptoms.

When we performed this analysis for the 20 genes with a significant (P<0.0016) association in the replication study, we found that for 14 genes the allergy protective allele was associated with increased gene expression (**Table E5** in this article's **Online Repository).** This includes, for example, the long non-coding RNA *RP11-534L20.5*, for which information from six different tissues (including blood, lung and skin) indicates that increased gene expression has a protective effect on disease risk. On the other hand, for four genes (*AC004893.11*, *ATXN2L*, *NSMCE1*, *RP11-24N18.1*) the allergy protective allele was associated with decreased gene expression, while for two genes eQTL directional effects were conflicting between studies (*IPCEF1*, *SULT1A1*).

Association analyses contrasting individuals suffering from a single allergic disease

The multiple-disease case-control phenotype analysed in our GWAS[6] maximizes power to identify risk variants that are shared between asthma, hay fever and eczema[26]. As such, we expected the 11 novel risk loci identified above to have comparable effects on the three individual diseases. To address this possibility, we tested the association between the 20 genes in those 11 loci and three phenotypes that compared three non-overlapping groups of adults who suffered from a single allergic disease: (1) asthma only cases (n=12,268) versus hay fever only cases (n=33,305); (2) asthma only cases versus eczema only cases (n=6,276); and (3) hay fever only cases versus eczema only cases. After correcting for the number of tests performed (P<0.05/(20 genes x 3 phenotypes)=0.0008), no single gene had a

311 significant association in the asthma versus hay fever, asthma versus eczema or hay fever versus
312 eczema analyses (**Table E6**). Even at a nominal $P<0.05$, which does not correct for multiple testing,
313 only three genes (in two loci) had a significant association in these analyses: *OR10J5*, *RP11-264B17*.4
314 and *LAT*. These results indicate that most (if not all) of the 11 novel risk loci identified in this study do
315 not have differential effects on the three individual diseases.
316

DISCUSSION

By analyzing results from our previous GWAS [6] with a gene-based test of association, followed by replication of top findings in an independent sample, we identified 11 loci that contain previously unrecognized genetic risk variants for allergic disease. Results from case-only association analyses indicate that these loci have similar effects on asthma, hay fever and eczema risk.

The 11 novel risk loci were not reported in the original GWAS because they did not contain any single SNP associated with disease risk at a significance threshold that accounted for the number of SNPs tested. We were able to identify these loci in the present study for two main reasons. First, by testing individual genes rather than SNPs, the multiple testing burden was greatly reduced; this translated into a less stringent threshold required to declare genome-wide significance (*P*-value of 2.5×10^{-6} instead of 3×10^{-8}), which increases power. Second, most of these loci (7 out of 11) contain genes for which multiple independent eQTL were individually associated with disease risk. Under this scenario, the gene-based test used improves power over the alternative approach of testing individual eQTL separately. Overall, our results support the use of gene-based eQTL-centric approaches to identify novel risk loci for human diseases and traits, as reported by others [10, 27].

Our results point to 19 genes as being the likely candidates underlying the association between the 11 new loci and allergic disease risk. We stress, however, that functional studies are now required to confirm that the expression of these genes (1) is determined by (not simply associated with) the eQTL included in the respective gene-based test; and (2) influences disease pathophysiology. In other words, for given locus, we cannot exclude the possibility that a gene with a significant association is not causally-related to disease pathophysiology. Instead, for example, it might simply share eQTL with a nearby causal gene that was not identified in our analysis. We highlight one possible example of this. We observed a genome-wide significant association with a single gene in the 1q23.2 locus: the olfactory receptor gene *OR10J5* ($P=2.3 \times 10^{-6}$). This gene is thought to regulate lipid accumulation [28],

and so could plausibly contribute to the pathophysiology of allergic disease. However, there were four additional genes within 1 Mb of *OR10J5* that had a statistically (albeit not genome-wide) significant gene-based association in our discovery analysis: *FCER1A* (P=1.2x10^{-5}), *DARC* (P=0.0005), *IFI16* (P=0.009) and *PYHIN1* (P=0.03), the latter reported previously to contain asthma risk variants in populations of African ancestry [29]. The association with the first two replicated in the independent UKB sample: P=4.8x10^{-7} and P=0.0001, respectively. Therefore, it is possible that *FCER1A* and/or *DARC* represent causal gene(s) underlying the observed association at this locus. Both are biologically plausible candidate allergy risk genes, encoding respectively the alpha subunit of the high-affinity IgE receptor and an atypical chemokine receptor [30]. Why did our analysis identify *OR10J5* and not, for example, *FCER1A*? The gene-based test for *OR10J5* included three eQTL (see **Table E2** in this article's **Online Repository)**, all individually associated (P<0.05) with disease risk; the strongest disease association was observed for eQTL rs2427838 (single-SNP P=1.0x10^{-5}). On the other hand, the gene-based test for *FCER1A* included these same three eQTL (*i.e.* these eQTL were shared between *OR10J5* and *FCER1A*) plus an additional six independent eQTL, none of which were individually associated with disease risk. Because the additional eQTL tested for *FCER1A* were not associated with disease risk, the resulting gene-based association was weaker when compared to *OR10J5*. An interesting question is why some but not all eQTL of *FCER1A* (and other genes) are associated with disease risk; it could be, for example, that disease associations are restricted to eQTL that influence gene expression in a specific subset of immune cells, or to eQTL that influence multiple relevant genes. Future studies that address this question are warranted.

With the caveat in mind that significant gene-based associations pinpoint causal risk loci but not necessarily the right causal gene(s), we note that nine of the 19 genes located in novel risk loci encode proteins with a known function that is directly relevant to allergic disease (**Table 2**). For example, *FOSL2* is involved in B cell and epidermal differentiation [31, 32]. Furthermore, it has a critical yet

complex role in Th17 differentiation and function [33]: on the one hand, it represses Th17 signature genes (e.g. *IL17A*), but on the other hand, it promotes the expression of genes that drive Th17 maintenance and survival (e.g. *IL6R*) [33]. When we compared the directional effect of *FOSL2* eQTL between disease risk and gene expression, we found that the allele associated with reduced disease risk was associated with increased gene expression. These genetic findings suggest that increased *FOSL2* expression results in attenuated allergic inflammatory responses. Consistent with this possibility, deletion of a *FOSL2* repressor in mice decreased the capacity of CD4+ T cells to develop into the pro-inflammatory T follicular helper cell lineage following immunization with ovalbumin, viral infection, or in the context of low-grade chronic inflammation[34]. Based on these observations, we suggest that therapeutic strategies that increase *FOSL2* expression should be considered for the treatment of allergic diseases.

Another example of a gene identified in our analysis and previously implicated in the pathophysiology of allergic disease was *IL27*. This was one of eight genes identified on chromosome 16p11.2, a region that overlaps a large (~0.45 Mb) and common (49% frequency in Europeans) genomic inversion previously reported to be associated with the joint occurrence of asthma and obesity [35]. Of the two eQTL included in the gene-based test for *IL27*, one (rs7191548) is in high LD (r^2=0.72) with a SNP that tags the inversion (rs4788101), suggesting that the association observed with *IL27* in our study is partly explained by that large structural variant. IL-27 has been shown to suppress Th2 responses [36, 37], and so has been suggested as a potential new therapy for asthma. However, to our knowledge, no clinical trials have been performed to test this possibility. Our observation that, for both *IL27* eQTL tested, the allele associated with reduced disease risk was associated with increased gene expression in blood provides further support for an anti-inflammatory effect of IL-27 for the treatment of allergic conditions.

Lastly, we identified four significant gene-based associations with non-coding RNAs of unknown function. Of these, the lncRNA *RP11-534L20.5* is of particular interest, as it is located in

close proximity (8kb) to *IKBKE*, a regulator of the NF-kappaB pathway that plays a role in immune-related mechanisms [38, 39]. Using data from release v7 of the GTEx project[40], we found a highly significant positive correlation in gene expression between *RP11-534L20.5* and *IKBKE* in the skin (P=10^{-11}), with a weaker but consistent effect in blood and lung (not shown). Such an association could arise, for example, if *RP11-534L20.5* regulates the expression of *IKBKE* or if both transcripts share a regulatory element. Consistent with the latter hypothesis, the 5' end of *RP11-534L20.5* overlaps a peak of H3K27 acetylation (a mark for active enhancers) in multiple cell lines and physically interacts with the *IKBKE* promoter in a B-cell line [41]. Further studies are warranted to investigate the function of RP11-534L20.5, as well as the other non-coding RNAs identified in our analysis.

Two additional caveats should be considered when interpreting results from our study. First, our original GWAS [6] included only individuals of European ancestry, and so it is unclear if the risk loci identified in our current study extend to individuals of other ancestries. Second, the association analyses performed to compare individuals suffering from a single allergic disease were based on a relatively small sample size, and so it is possible that the lack of significant associations reflects the lower power of these analyses.

In conclusion, we identified significant and reproducible gene-based associations with 19 genes located in 11 loci not previously reported in GWAS of allergic disease. Our genetic findings suggest that drugs that target these genes might have an increased probability of success if prioritised for clinical development [42]. Our results further demonstrate the utility of applying gene-based tests of association to results from existing GWAS.

ACKNOWLEDGMENTS

T This research has been conducted using the UK Biobank Resource under Application Number 10074. MARF was supported by a Senior Research Fellowship (APP1124501) from the National Health and Medical Research Council (NHMRC) of Australia. JDH was supported by NIH postdoctoral training grant CA112355. LP was funded by a UK MRC fellowship award (MR/J012165/1) and works in a unit funded by the UK MRC (MC_UU_12013). Detailed acknowledgments are provided in the Online Repository.

REFERENCES

1. Fagnani C, Annesi-Maesano I, Brescianini S, D'Ippolito C, Medda E, Nistico L, et al. Heritability and shared genetic effects of asthma and hay fever: an Italian study of young twins. Twin Res Hum Genet 2008; 11:121-31.
2. Duffy DL, Martin NG, Battistutta D, Hopper JL, Mathews JD. Genetics of asthma and hay fever in Australian twins. Am Rev Respir Dis 1990; 142:1351-8.
3. Thomsen SF, Ulrik CS, Kyvik KO, Skadhauge LR, Steffensen I, Backer V. Findings on the atopic triad from a Danish twin registry. Int J Tuberc Lung Dis 2006; 10:1268-72.
4. van Beijsterveldt CE, Boomsma DI. Genetics of parentally reported asthma, eczema and rhinitis in 5-yr-old twins. Eur Respir J 2007; 29:516-21.
5. Ullemar V, Magnusson PK, Lundholm C, Zettergren A, Melen E, Lichtenstein P, et al. Heritability and confirmation of genetic association studies for childhood asthma in twins. Allergy 2016; 71:230-8.
6. Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. Nat Genet 2017; 49:1752-7.
7. Fadista J, Manning AK, Florez JC, Groop L. The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. Eur J Hum Genet 2016; 24:1202-5.
8. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, et al. A versatile gene-based test for genome-wide association studies. Am J Hum Genet 2010; 87:139-45.
9. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 2015; 11:e1004219.
10. Bakshi A, Zhu Z, Vinkhuyzen AA, Hill WD, McRae AF, Visscher PM, et al. Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. Sci Rep 2016; 6:32894.
11. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 2010; 6:e1000888.
12. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 2015; 47:1091-8.
13. Ferreira MA, Jansen R, Willemsen G, Penninx B, Bain LM, Vicente CT, et al. Gene-based analysis of regulatory variants identifies 4 putative novel asthma risk genes related to nucleotide synthesis and signaling. J Allergy Clin Immunol 2017; 139:1148-57.
14. Barbeira A, Dickinson SP, Torres JM, Bonazzola R, Zheng J, Torstenson ES, et al. Integrating tissue specific mechanisms into GWAS summary results. bioRxiv 2017.
15. Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot

analysis: multitissue gene regulation in humans. Science 2015; 348:648-60.
16. Odhams CA, Cunninghame Graham DS, Vyse TJ. Profiling RNA-Seq at multiple resolutions markedly increases the number of causal eQTLs in autoimmune disease. PLoS Genet 2017; 13:e1007071.
17. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. Genome-wide genetic data on ~500,000 UK Biobank participants. bioRxiv 2017.
18. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015; 47:291-5.
19. Yang J, Ferreira T, Morris AP, Medland SE, Madden PA, Heath AC, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 2012; 44:369-75, S1-3.
20. Davis JR, Fresard L, Knowles DA, Pala M, Bustamante CD, Battle A, et al. An Efficient Multiple-Testing Adjustment for eQTL Studies that Accounts for Linkage Disequilibrium between Variants. Am J Hum Genet 2016; 98:216-24.
21. Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, et al. Transcriptome genetics using second generation sequencing in a Caucasian population. Nature 2010; 464:773-7.
22. Lappalainen T, Sammeth M, Friedlander MR, t Hoen PA, Monlong J, Rivas MA, et al. Transcriptome and genome sequencing uncovers functional variation in humans. Nature 2013; 501:506-11.
23. Pierce BL, Tong L, Chen LS, Rahaman R, Argos M, Jasmine F, et al. Mediation analysis demonstrates that trans-eQTLs are often explained by cis-mediation: a genome-wide analysis among 1,800 South Asians. PLoS Genet 2014; 10:e1004818.
24. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015; 4:7.
25. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015.
26. Ferreira MA. Improving the power to detect risk variants for allergic disease by defining case-control status based on both asthma and hay fever. Twin Res Hum Genet 2014; 17:505-11.
27. Xu Z, Wu C, Wei P, Pan W. A Powerful Framework for Integrating eQTL and GWAS Summary Data. Genetics 2017.
28. Tong T, Ryu SE, Min Y, de March CA, Bushdid C, Golebiowski J, et al. Olfactory receptor 10J5 responding to alpha-cedrene regulates hepatic steatosis via the cAMP-PKA pathway. Sci Rep 2017; 7:9471.

29. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. Nat Genet 2011; 43:887-92.
30. Graham GJ, Locati M, Mantovani A, Rot A, Thelen M. The biochemistry and biology of the atypical chemokine receptors. Immunol Lett 2012; 145:30-8.
31. Ubieta K, Garcia M, Grotsch B, Uebe S, Weber GF, Stein M, et al. Fra-2 regulates B cell development by enhancing IRF4 and Foxo1 transcription. J Exp Med 2017; 214:2059-71.
32. Wurm S, Zhang J, Guinea-Viniegra J, Garcia F, Munoz J, Bakiri L, et al. Terminal epidermal differentiation is regulated by the interaction of Fra-2/AP-1 with Ezh2 and ERK1/2. Genes Dev 2015; 29:144-56.
33. Ciofani M, Madar A, Galan C, Sellars M, Mace K, Pauli F, et al. A validated regulatory network for Th17 cell specification. Cell 2012; 151:289-303.
34. Hu R, Kagele DA, Huffaker TB, Runtsch MC, Alexander M, Liu J, et al. miR-155 promotes T follicular helper cell accumulation during chronic, low-grade inflammation. Immunity 2014; 41:605-19.
35. Gonzalez JR, Caceres A, Esko T, Cusco I, Puig M, Esnaola M, et al. A common 16p11.2 inversion underlies the joint susceptibility to asthma and obesity. Am J Hum Genet 2014; 94:361-72.
36. Jirmo AC, Daluege K, Happle C, Albrecht M, Dittrich AM, Busse M, et al. IL-27 Is Essential for Suppression of Experimental Allergic Asthma by the TLR7/8 Agonist R848 (Resiquimod). J Immunol 2016; 197:4219-27.
37. Yoshimoto T, Yoshimoto T, Yasuda K, Mizuguchi J, Nakanishi K. IL-27 suppresses Th2 cell development and Th2 cytokines production from polarized Th2 cells: a novel therapeutic way for Th2-mediated allergic inflammation. J Immunol 2007; 179:4415-23.
38. Alves BN, Tsui R, Almaden J, Shokhirev MN, Davis-Turak J, Fujimoto J, et al. IkappaBepsilon is a key regulator of B cell expansion by providing negative feedback on cRel and RelA in a stimulus-specific manner. J Immunol 2014; 192:3121-32.
39. Ng SL, Friedman BA, Schmid S, Gertz J, Myers RM, Tenoever BR, et al. IkappaB kinase epsilon (IKK(epsilon)) regulates the balance between type I and type II interferon responses. Proc Natl Acad Sci U S A 2011; 108:21170-5.
40. Consortium GT, Laboratory DA, Coordinating Center -Analysis Working G, Statistical Methods groups-Analysis Working G, Enhancing Gg, Fund NIHC, et al. Genetic effects on gene expression across human tissues. Nature 2017; 550:204-13.
41. Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L, et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat Genet 2015; 47:598-606.
42. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, et al. The support of human genetic evidence for approved drug indications. Nat Genet

2015; 47:856-60.
43. Pickrell JK, Berisa T, Liu JZ, Segurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet 2016; 48:709-17.
44. Wrann CD, Eguchi J, Bozec A, Xu Z, Mikkelsen T, Gimble J, et al. FOSL2 promotes leptin gene expression in human and mouse adipocytes. J Clin Invest 2012; 122:1010-21.
45. Guo Z, Kong Q, Liu C, Zhang S, Zou L, Yan F, et al. DCAF1 controls T-cell function via p53-dependent and -independent mechanisms. Nat Commun 2016; 7:10307.
46. Kassmeier MD, Mondal K, Palmer VL, Raval P, Kumar S, Perry GA, et al. VprBP binds full-length RAG1 and is required for B-cell development and V(D)J recombination fidelity. EMBO J 2012; 31:945-58.
47. Sadewasser A, Paki K, Eichelbaum K, Bogdanow B, Saenger S, Budt M, et al. Quantitative Proteomic Approach Identifies Vpr Binding Protein as Novel Host Factor Supporting Influenza A Virus Infections in Human Cells. Mol Cell Proteomics 2017; 16:728-42.
48. Venkateswarlu K. Interaction protein for cytohesin exchange factors 1 (IPCEF1) binds cytohesin 2 and modifies its activity. J Biol Chem 2003; 278:43460-9.
49. Zhu W, London NR, Gibson CC, Davis CT, Tong Z, Sorensen LK, et al. Interleukin receptor activates a MYD88-ARNO-ARF6 cascade to disrupt vascular stability. Nature 2012; 492:252-5.
50. Oh SJ, Santy LC. Differential effects of cytohesins 2 and 3 on beta1 integrin recycling. J Biol Chem 2010; 285:14610-6.
51. Thedieck K, Polak P, Kim ML, Molle KD, Cohen A, Jeno P, et al. PRAS40 and PRR5-like protein are new mTOR interactors that regulate apoptosis. PLoS One 2007; 2:e1217.
52. Ellson CD, Davidson K, Ferguson GJ, O'Connor R, Stephens LR, Hawkins PT. Neutrophils from p40phox-/- mice exhibit severe defects in NADPH oxidase regulation and oxidant-dependent bacterial killing. J Exp Med 2006; 203:1927-37.
53. Crotzer VL, Matute JD, Arias AA, Zhao H, Quilliam LA, Dinauer MC, et al. Cutting edge: NADPH oxidase modulates MHC class II antigen presentation by B cells. J Immunol 2012; 189:3800-4.
54. Brown ML, Ramprasad MP, Umeda PK, Tanaka A, Kobayashi Y, Watanabe T, et al. A macrophage receptor for apolipoprotein B48: cloning, expression, and atherosclerosis. Proc Natl Acad Sci U S A 2000; 97:7488-93.
55. Do J, Kim D, Kim S, Valentin-Torres A, Dvorina N, Jang E, et al. Treg-specific IL-27Ralpha deletion uncovers a key role for IL-27 in Treg function to control autoimmunity. Proc Natl Acad Sci U S A 2017; 114:10190-5.
56. Muallem G, Wagage S, Sun Y, DeLong JH, Valenzuela A, Christian DA, et al. IL-

27 Limits Type 2 Immunopathology Following Parainfluenza Virus Infection. PLoS Pathog 2017; 13:e1006173.

57. Yang B, Suwanpradid J, Sanchez-Lagunes R, Choi HW, Hoang P, Wang D, et al. IL-27 Facilitates Skin Wound Healing through Induction of Epidermal Proliferation and Host Defense. J Invest Dermatol 2017; 137:1166-75.

58. Meunier C, Bordereaux D, Porteu F, Gisselbrecht S, Chretien S, Courtois G. Cloning and characterization of a family of proteins associated with Mpl. J Biol Chem 2002; 277:9139-47.

59. Bacchelli C, Moretti FA, Carmo M, Adams S, Stanescu HC, Pearce K, et al. Mutations in linker for activation of T cells (LAT) lead to a novel form of severe combined immunodeficiency. J Allergy Clin Immunol 2017; 139:634-42 e5.

60. Aguado E, Richelme S, Nunez-Cruz S, Miazek A, Mura AM, Richelme M, et al. Induction of T helper type 2 immunity by a point mutation in the LAT adaptor. Science 2002; 296:2036-40.

61. Zhang W, Sloan-Lancaster J, Kitchen J, Trible RP, Samelson LE. LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor to cellular activation. Cell 1998; 92:83-92.

605 **FIGURE LEGENDS**

606 **Figure 1. Outline of analytical approach used.**

607 **Figure 2. Summary of association results between allergic disease status and single SNPs or**
608 **individual genes.** (A) Single-SNP association results published by Ferreira et al. [6]. (B) Single-SNP
609 association results obtained after adjusting the results of Ferreira et al. for the effect of the 136 genome-
610 wide significant associations reported in that study. Variants in four loci (in *CADM3*, *SLC39A8*,
611 *LRRC43* and between *KLF5* and *KLF12*) were genome-wide significant (*i.e.* $P<3x10^{-8}$) in the adjusted
612 but not in the original GWAS results, and so represent secondary single-SNP association signals in loci
613 identified in Ferreira et al. [6]. (C) Gene-based association results obtained for 19,432 genes after
614 applying the EUGENE approach to results from the adjusted GWAS. Genes with a gene-based
615 association $P<2.5x10^{-6}$ are listed, with font color reflecting the evidence for association with those
616 genes in the UK Biobank replication study: green - $P<0.0016$ (*i.e.* significant after correcting for 30
617 genes tested); blue - $0.0016<P<0.05$ (*i.e.* nominally significant); and black - $P>0.05$ (*i.e.* not
618 significant).

619 **Figure 3. Comparison of results obtained with single-SNP and gene-based analyses for the 30**
620 **genes with a significant gene-based P-value (*i.e.* $P<2.5x10^{-6}$) in the discovery study.** For each gene,
621 the *x*-axis shows the statistical evidence for association (-log10*P*-value) obtained with the eQTL most
622 strongly associated with disease risk. The *y*-axis shows the evidence for association (-log10*P*-value)
623 obtained for each gene with the EUGENE gene-based approach. Gene names are shown for eight genes
624 for which the latter was at least one order of magnitude more significant than the former. The color of
625 each circle indicates the number of independent eQTL for that gene that were individually associated
626 with disease risk at a $P<0.05$: black (one associated eQTL), green (two), yellow (three) and red (three
627 or more).

TABLES

Table 1. Association results for 30 genes with a gene-based association $P<2.5x10^{-6}$ in the discovery analysis.

Locus	Gene	Gene location			eQTL with strongest association in GWAS			Gene-based analysis: discovery (n=360,838)				Gene-based analysis: replication (n=233,898)		
		Chr	Start		SNP	GWAS P-value		N eQTLs tested	N eQTLs with P<0.05	EUGENE P-value		N eQTLs tested	N eQTLs with P<0.05	EUGENE P-value
1	*CASZ1**	1	10696661		rs12045923	4.9E-07		2	1	1.9E-06		2	1	*2.0E-02*
2	*OR10J5*	1	159504793		rs2427837	1.0E-05		3	3	2.3E-06		3	3	**6.3E-07**
3	*RP11-534L20.5*	1	206677281		rs11117858	8.1E-06		12	5	1.3E-06		11	3	**1.1E-03**
4	*FOSL2*	2	28615315		rs7562	8.8E-07		1	1	8.8E-07		1	1	**2.6E-04**
5	*RBM15B*	3	51428731		rs73078636	1.4E-07		1	1	1.4E-07		1	1	**4.0E-04**
	VPRBP	3	51433298		rs73078636	1.4E-07		1	1	1.4E-07		1	1	**4.0E-04**
6	*TICAM2*	5	114914339		rs17137937	2.9E-06		3	2	1.7E-06		3	1	*5.7E-03*
7	*NUP43*	6	150045451		rs6909158	7.5E-06		22	6	2.1E-06		22	1	4.5E-01
8	*IPCEF1*	6	154475631		rs9397706	4.7E-07		4	3	4.3E-07		4	3	**4.1E-07**
9	*AC004893.11*	7	98610788		rs4236540	8.1E-07		1	1	8.1E-07		1	1	**4.3E-04**
10	*ABO**	9	136125788		rs550057	2.8E-07		14	5	1.7E-07		14	6	**3.8E-10**
11	*PTPLA*	10	17631958		rs7092926	6.8E-07		4	4	8.2E-07		4	0	2.1E-01
	TMEM236	10	17794251		rs7092926	6.8E-07		3	3	1.9E-06		3	0	3.6E-01
12	*PRR5L*	11	36317838		rs7925585	2.8E-06		6	2	9.7E-07		6	2	**1.4E-06**
13	*NSMCE1*	16	27236312		rs4523932	5.2E-06		2	2	6.6E-07		2	2	**1.3E-06**
14	*APOBR*	16	28505970		rs151233	3.2E-06		2	2	1.6E-08		2	2	**3.8E-05**
	IL27	16	28510683		rs231970	4.8E-06		2	2	2.6E-08		2	2	**1.6E-04**
	SULT1A1	16	28616903		rs75539558	1.6E-06		8	3	2.0E-07		8	4	**1.4E-05**
	ATXN2L	16	28834356		rs8056890	1.7E-06		1	1	1.7E-06		1	1	**1.7E-04**
	RP11-24N18.1	16	28841933		rs8056890	1.7E-06		1	1	1.7E-06		1	1	**1.7E-04**
	SPNS1	16	28985542		rs2726040	4.8E-06		15	6	3.5E-09		14	5	**7.4E-04**
	RP11-264B17.4	16	28986294		rs8045689	6.3E-06		3	2	1.1E-06		3	2	**1.5E-03**
	LAT	16	28996147		rs8045689	6.3E-06		3	2	1.1E-06		3	3	**2.4E-04**
15	*MYO1C*	17	1367392		rs56157500	9.2E-05		4	3	1.5E-06		4	1	5.6E-02
16	*ICAM1*	19	10381511		rs1799969	1.8E-05		12	7	5.1E-07		12	0	4.1E-01
	CTD-2369P2.8	19	10396477		rs1799969	1.8E-05		12	7	5.1E-07		12	0	4.1E-01
	ICAM4	19	10397643		rs1799969	1.8E-05		12	7	6.1E-07		11	0	4.9E-01
17	*FPR1*	19	52248425		rs7254019	1.8E-05		14	7	1.5E-08		14	3	*3.2E-03*
18	*PVALB*	22	37196728		rs4821544	3.6E-07		5	2	1.3E-07		5	4	**5.9E-05**

	NCF4	22	37257030		rs4821544	3.6E-07		5	2	4.8E-08		5	3	**1.1E-04**

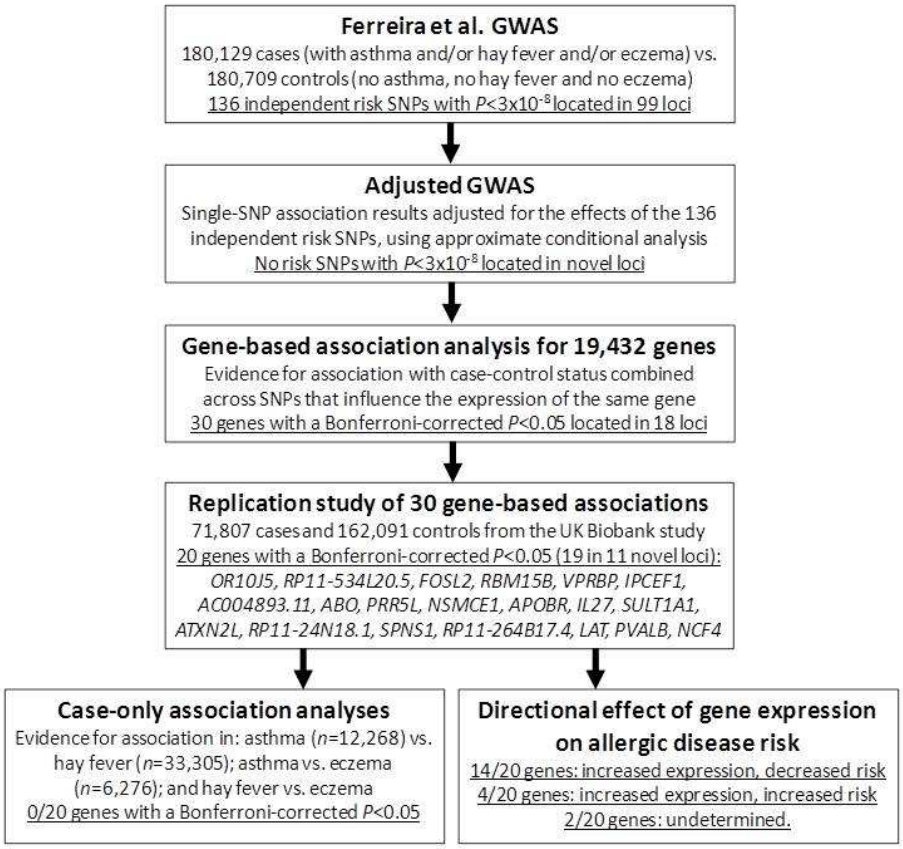
* Located in a locus identified in a previous GWAS of allergic disease [43]

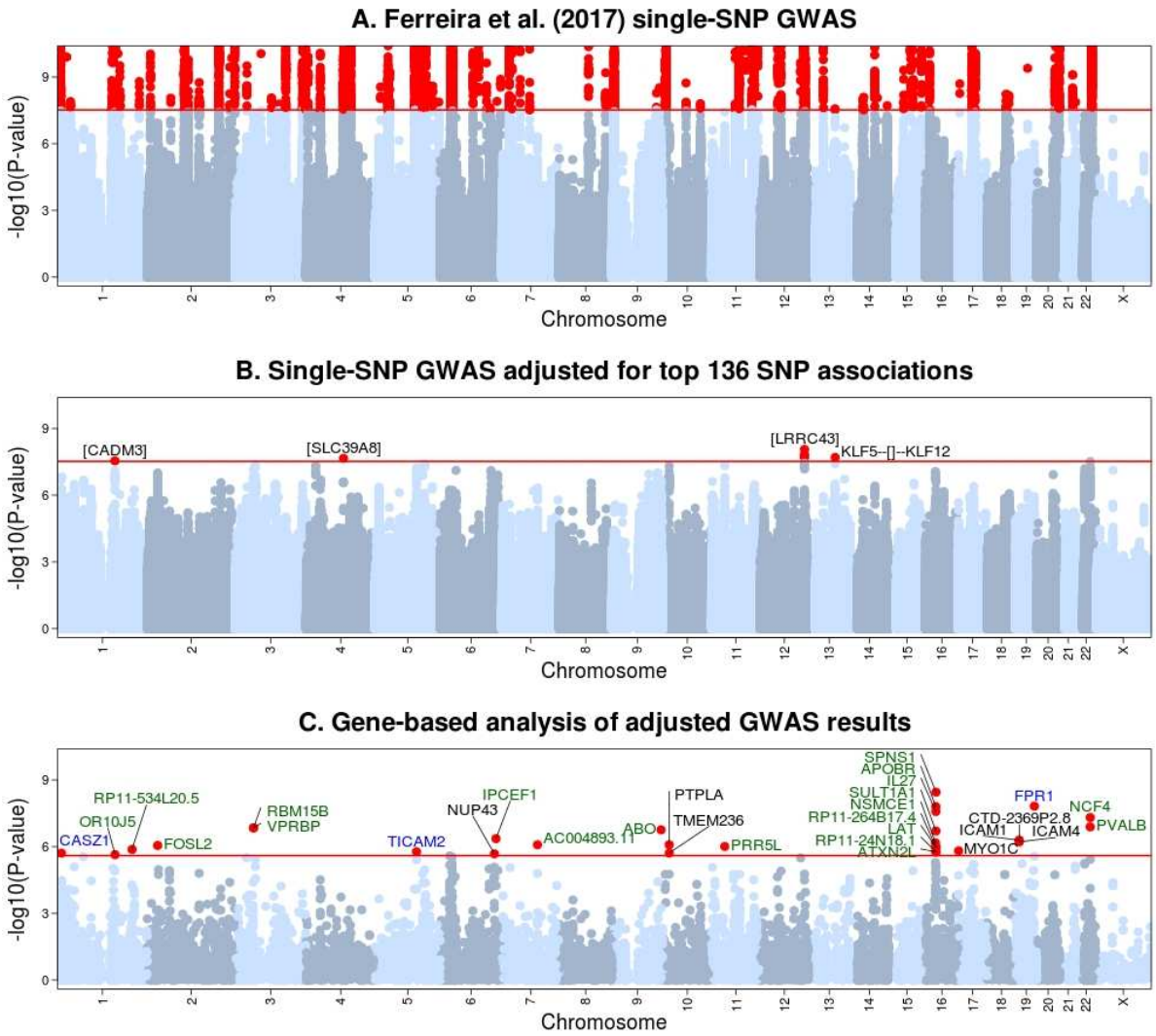
The replication P-value is in bold if significant after correcting for the 30 genes tested (*i.e.* $P<0.05/30=0.0016$) and in italic if $0.0016<P<0.05$ (i.e. nominally significant).

634 **Table 2. Genes with a known function that is relevant to the pathophysiology of allergic disease.**

Gene	Relevant function	References
FOSL2	B cell and epidermal differentiation, Th17 and adipocyte function	31-33, 44
VPRBP	T cell function, B cell development, viral replication	45-47
IPCEF1	Binds cytohesin 2, which is involved in IL-1beta signaling and cell adhesion	48-50
PRR5L	mTOR interactor that regulates cell death	51
NCF4	Phagocyte oxidative burst, antigen presentation	52, 53
APOBR	Macrophage receptor for apolipoprotein B48	54
IL27	Regulation of Th1 and Th2 reponses, Treg function, epithelial cell proliferation	37, 55-57
ATXN2L	Cytokine signaling	58
LAT	T cell development and activation	59-61

635





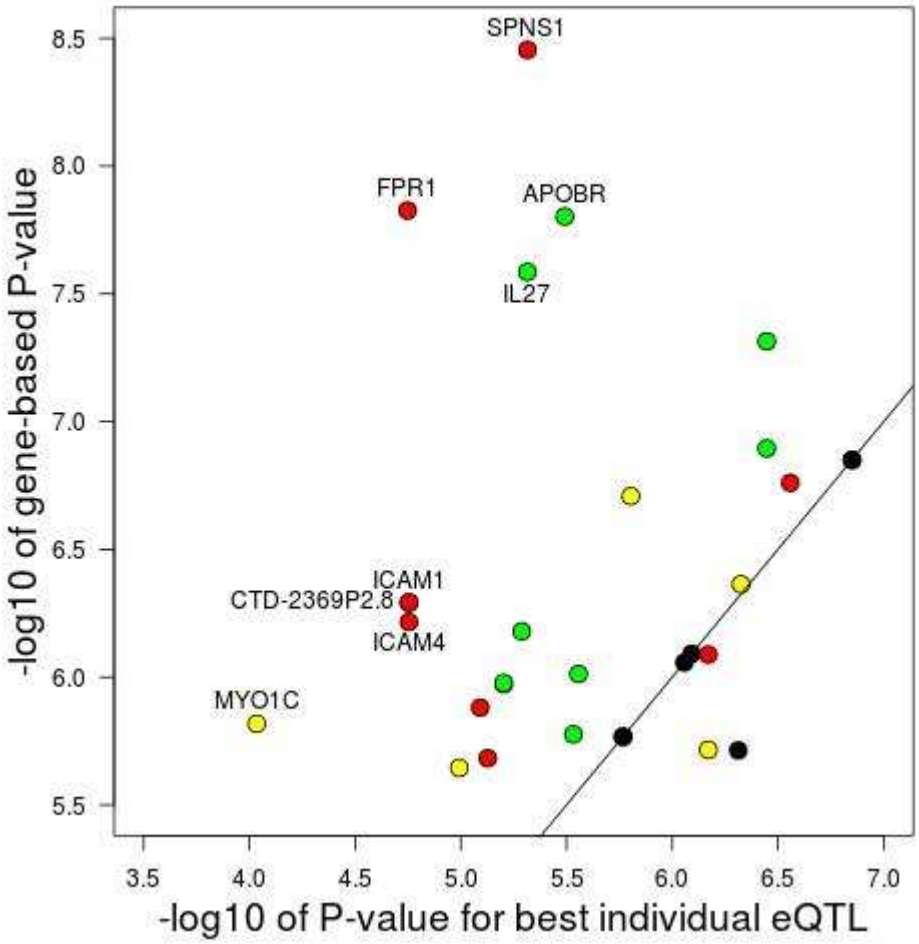


Table E1. Thirteen studies that contributed to the GWAS of allergic disease reported by Ferreira et al. 2017.

Study	Study-type	Original case ascertainment	N total	N controls	Number of cases								N cases with co-morbidity	Age (mean, range)	% Females	Mean age-of-onset (range)		
					Total	With only a single disease			With two diseases			With all three diseases	Unclassifiable#					
						A+H-E-	A-H+E-	A-H-E+	A+H+E-	A+H-E+	A-H+E+	A+H+E+				Asthma	Hay fever	Eczema
UKBiobank	Population-based	NA	138354	96108	42246	8769	4838	1538	1661	596	439	364	24041	3060	56.7 (39-73)	53	30.3 (1-73)	24.4 (1-70)	25.6 (1-69)
23andMe	Population-based	NA	118269	34934	83335	2048	26125	3648	8574	365	4193	3610	34772	16742	49.5 (1-114)	48	20.8 (0-93)	NA	23.6 (0-96)
GERA	Population-based	NA	51218	15999	35219	4439	17401	2297	7836	416	1839	991	0	11082	62.3 (18-90)	59	NA	NA	NA
CATSS	Population-based	NA	11068	7488	3580	1600	380	681	387	240	98	194	0	919	9.8 (9-23)	49	4.2 (0-23)	5.0 (0-11)	1.5 (0-11)
NTR	Population-based	NA	10242	7919	2323	685	837	219	421	64	48	49	0	582	40.1 (4-94)	64	NA	NA	NA
LifeLines	Population-based	NA	8560	4837	3723	190	1718	897	245	58	468	113	34	884	46.2 (18-88)	58	17.8 (0-66)	23.4 (0-88)	NA
TWINGENE	Population-based	NA	5517	3762	1755	1007	418	79	219	21	5	6	0	251	58.3 (41-93)	51	54.1 (0-93)	24.9 (0-72)	1.5 (0-45)
ALSPAC$	Population-based	NA	4964	2330	2634	373	620	374	224	139	187	187	530	737	A/E: 10.8 (10-13); H: 13.9 (13-16)	49	NA	NA	NA
SALTY	Population-based	NA	4062	2761	1301	583	365	79	226	27	14	7	0	274	49.8 (41-72)	49	41.6 (0-72)	20.3 (0-52)	8.9 (0-46)
AAGC	Selected case-control	Asthma	2435	460	1975	393	124	32	701	156	14	555	0	1426	35.1 (3-89)	56	15.5 (0-75)	NA	6.7 (0-40)
GENEVA	Selected case-control	Eczema	2633	1274	1359	41	154	383	29	65	309	289	89	692	43.9 (0-85)	56	NA	NA	11.7 (0-85)
GENUFAD-SHIP-1	Selected case-control	Eczema	1781	1364	417	0	0	323	0	46	26	22	0	94	Cases: 3.9 (1-34); Controls: 50.0 (20-81)	Cases: 39; Controls: 50	7 (1-31)	8 (2-31)	1 (0-2)
GENUFAD-SHIP-2	Selected case-control	Eczema	1735	1473	262	0	0	169	0	35	27	31	0	93	Cases: 8.3 (1-26); Controls: 50.0 (20-81)	Cases: 54; Controls: 50	11 (2-24)	13 (1-26)	1 (0-2)
Total			360838	180709	180129	20128	52980	10719	20523	2228	7667	6418	59466	36836					

Individuals with missing information for at least one disease could not be classified into one of the seven three-disease groups.
$ For ALSPAC, information from different surveys was used to define asthma (A) and eczema (E) when compared to those used to define hay fever. For this reason, age of participants is reported separately for A/E and H.

Table E2. Genome-wide association studies of gene expression levels queried to identify expression quantitative trait loci (i.e. eQTL)

First author	PMID	Tissue/cell type	Sample size	cis effects	
				N genes	N associations P<2.3x10-9
Andiappan	26259071	NEUTROPHILS	114	310	3776
Barreiro	22233810	DENDRITIC-NotInfected	65	36	36
		DENDRITIC-TBInfected	65	216	216
Battle	24092820	WHOLE-BLOOD	922	7792	7793
		WHOLE-BLOOD-ase	922	310	537
		WHOLE-BLOOD-splice	922	532	818
Brumpton	27155841	LCLS	356	27	2139
Caliskan	25874939	PBMCS-baseline	98	173	173
		PBMCS-rhinovirus	98	169	169
		PBMCS-rhinovirus-reQTL	98	12	12
Davenport	26917434	LEUCOCYTES	265	879	7699
Dimas	19644074	LCLS	75	97	121
		TCELLS	75	85	107
		FIBROBLASTS	75	102	136
Di Narzo	27336838	WHOLE-BLOOD	149	1195	2220
Ding	21129726	LESIONAL-SKIN	57	15	295
		NORMAL-SKIN	53	19	350
		UNINVOLVED-SKIN	53	12	160
Dixon	17873877	LCLS	400	634	6735
Fairfax 2012	22446964	BCELLS	283	236	1723
		MONOCYTES	283	549	4051
Fairfax 2014	24604202	MONOCYTES-IFN	367	2048	21366
		MONOCYTES-LPS2	261	686	4658
		MONOCYTES-LPS24	322	1412	10740
		MONOCYTES-NAIVE	414	2024	24676
Fehrmann	21829388	WHOLE-BLOOD	1469	2944	23614
Ferraro	24610777	Tconv	65	26	118
		Tregs	65	24	105
Franco	23878721	WHOLE-BLOOD-Influenza	247	65	329
Lappalainen	24037378	LCLS	373	3525	1397237
Grundberg	22941192	LCLS	856	1550	68025
		SKIN	856	879	31390
GTEx	25954001	FIBROBLASTS	272	3910	368252
		LCLS	114	1062	88242
		LUNG	278	3148	335019

		SKIN	302	4026	613093
		SPLEEN	89	926	74864
		WHOLE-BLOOD	338	3331	343913
Hao	23209423	LUNG	1111	4177	6339
Huang	25951796	LCLs	368	2196	3091
		PBMCs	240	1939	2617
		SKIN	110	384	453
Jansen	Under Review	WHOLE-BLOOD	4896	4377	7460
Kasela	28248954	CD4TCELLS	293	1256	121508
		CD8TCELLS	293	947	83337
Kim	25327457	MONOCYTES-Baseline	137	681	5669
		MONOCYTES-Differential	137	62	377
		MONOCYTES-LPS	137	554	4802
Kukurba	27197214	WHOLE-BLOOD	922	90	8333
Lee	24604203	DENDRITIC-Baseline	528	82	82
		DENDRITIC-Flu	342	105	105
		DENDRITIC-Flu-delta	342	44	44
		DENDRITIC-IFNb	284	86	86
		DENDRITIC-IFNb-delta	284	23	24
		DENDRITIC-LPS	356	96	96
		DENDRITIC-LPS-delta	356	31	31
LloydJones	28065468	WHOLE-BLOOD	2765	5669	1373201
Luo	26102239	SMALL-AIRWAYS	105	152	174
Murphy	20833654	CD4-TCELLS	200	294	986
Naranbhai	26151758	NEUTROPHILS	101	493	547
Nedelec	27768889	MACROPHAGES-baseline	95	451	451
		MACROPHAGES-baseline-asQTL	95	634	951
		MACROPHAGES-listeria	95	422	422
		MACROPHAGES-listeria-asQTL	95	557	827
		MACROPHAGES-listeria-reQTL	95	93	93
		MACROPHAGES-salmonella	95	433	434
		MACROPHAGES-salmonella-asQTL	95	487	741
		MACROPHAGES-salmonella-reQTL	95	199	199
Peters	27015630	BCELLS	80	66	318
		CD4-TCELLS	121	396	2187
		CD8-TCELLS	108	277	1484
		MONOCYTES	124	564	3254
		NEUTROPHILS	121	341	1766

Table E2

Quach	27768888	MONOCYTES-baseline	100	467	469
		MONOCYTES-IAV	100	366	366
		MONOCYTES-LPS	100	464	465
		MONOCYTES-Pam3CSK4	100	536	537
		MONOCYTES-R848	100	478	479
Raj	24786080	CD4-TCELLS	407	1898	150019
		MONOCYTES	401	2494	196870
Walsh	27140173	WHOLE-BLOOD	377	3777	546913
Westra	24013639	WHOLE-BLOOD	5311	4285	343167
Yao	28285768	WHOLEBLOOD	5257	1681	93834
Ye	25214635	CD4-TCELLS-48h328	348	51	51
		CD4-TCELLS-48hTh17	348	45	45
		CD4-TCELLS-4h328	348	24	24
		CD4-TCELLS-4hIFNb	348	27	27
		CD4-TCELLS-UNST	348	19	19
Zeller	20502693	MONOCYTES	1490	2322	31353
Zhernakova	27918533	WHOLEBLOOD-exon-primary	2116	15028	88928
		WHOLEBLOOD-exonratio-primary	2116	6064	20222
		WHOLEBLOOD-gene-contextspecific	2116	14670	27186
		WHOLEBLOOD-gene-primary	2116	15158	3669632
		WHOLEBLOOD-polyAratio-primary	2116	1286	1843

Table E3. eQTL included in the gene based test for each of the 30 genes identified in the discovery study. Table E3

Locus	Gene	eQTLs included in gene-based test
1	*CASZ1*	rs12045923|rs284296
2	*OR10J5*	rs2427837|rs2494260|rs6699459
3	*RP11-534L20.5*	rs11117858|rs17433909|rs2297546|rs2336941|rs2987936|rs35785716|rs41299005|rs61816859|rs6666087|rs74882519|rs7538261|rs874718
4	*FOSL2*	rs7562
5	*RBM15B*	rs73078636
5	*VPRBP*	rs73078636
6	*TICAM2*	rs17137937|rs2546480|rs256938
7	*NUP43*	rs112165893|rs117067995|rs12523685|rs13205080|rs139394852|rs17733403|rs60328093|rs2151910|rs2281436|rs237004|rs2789503|rs4870058|rs4870139|rs62439806|rs
8	*IPCEF1*	rs12528985|rs1406055|rs7759388|rs9397706
9	*AC004893.11*	rs4236540
10	*ABO*	rs10901244|rs12379977|rs12683493|rs176694|rs28645493|rs3094377|rs35664240|rs4962050|rs550057|rs623361|rs62576042|rs626035|rs9411463|rs9411490
11	*PTPLA*	rs17141430|rs7092926|rs7094705|rs78704091
11	*TMEM236*	rs45607131|rs55999004|rs7092926
12	*PRR5L*	rs10836545|rs1123347|rs12270539|rs12785381|rs429034|rs7925585
13	*NSMCE1*	rs4523932|rs9788909
14	*SPNS1*	rs112918513|rs11569775|rs11643913|rs12448482|rs139456978|rs151230|rs2726040|rs62035317|rs67479058|rs7201546|rs7498329|rs7499778|rs75556002|rs8045689|rs8
14	*APOBR*	rs151233|rs27741
14	*IL27*	rs231970|rs7191548
14	*SULT1A1*	rs11540497|rs12935321|rs79476281|rs231970|rs4788115|rs4788119|rs75227850|rs75539558
14	*RP11-264B17.4*	rs12448482|rs67479058|rs8045689
14	*LAT*	rs4788115|rs6565261|rs8045689
14	*ATXN2L*	rs8056890
14	*RP11-24N18.1*	rs8056890
15	*MYO1C*	rs147003567|rs4790152|rs56157500|rs59343110
16	*CTD-2369P2.8*	rs10411056|rs1799969|rs2017213|rs2358581|rs2633973|rs281437|rs28378712|rs3177696|rs62130661|rs77123125|rs78064630|rs79337061
16	*ICAM1*	rs10411056|rs1799969|rs2017213|rs2358581|rs2633973|rs281437|rs28378712|rs3177696|rs62130661|rs77123125|rs78064630|rs79337061
16	*ICAM4*	rs10411056|rs1799969|rs2017213|rs2358581|rs2633973|rs281437|rs28378712|rs3177696|rs62130661|rs7252450|rs77123125|rs78064630
17	*FPR1*	rs117223215|rs12461801|rs16983230|rs17661285|rs17803207|rs1868943|rs62106945|rs6509571|rs7253284|rs7254019|rs73056872|rs79956121|rs8104640|rs8110040
18	*NCF4*	rs2072710|rs4821544|rs4821549|rs5756363|rs5756391
18	*PVALB*	rs1015775|rs4820250|rs4821544|rs730483|rs74971025

Table E3

;6918774|rs6922028|rs7762285|rs78765468|rs79782857|rs9379347|rs6909158|rs9942443

;056259

Table E4. Association results for each eQTL included in the gene-based test of each of the 30 genes identified in the discovery analysis.

Gene	EUGENE gene-based P-value	Single-SNP association results with allergic disease in the adjusted GWAS						Single-SNP association results with gene expression in the respective published eQTL study *				
		Proxy tested	P-value	Effect allele	Other allele	Effect allele frequency	Beta	eQTL	eQTL study and tissue	P-value	Effect allele	Effect
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	Zhernakova_WHOLEBLOOD-gene-primary	6.05E-142	T	-25.3655
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	Zhernakova_WHOLEBLOOD-gene-contextspecific	6.05E-142	T	-25.3655
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	Zhernakova_WHOLEBLOOD-exon-primary	6.54E-133	T	-24.5331
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	Zhernakova_WHOLEBLOOD-exonratio-primary	1.53E-29	T	-11.2868
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	Zhernakova_WHOLEBLOOD-exon-primary	1.04E-27	T	-10.9092
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	Zhernakova_WHOLEBLOOD-exon-primary	8.92E-27	T	-10.712
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	Zhernakova_WHOLEBLOOD-exon-primary	4.48E-26	T	-10.5617
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	Zhernakova_WHOLEBLOOD-exonratio-primary	8.08E-24	T	-10.0626
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	Zhernakova_WHOLEBLOOD-exonratio-primary	1.85E-23	T	-9.98077
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	Yao_WHOLEBLOOD	1.28E-11	C	0.0253396
ABO	1.74E-07	rs550057	2.77E-07	T	C	0.2806	0.0294	rs550057	GTE_WHOLEBLOOD	3.45E-10	T	-0.53797
ABO	1.74E-07	rs176694	1.15E-05	T	G	0.8895	-0.0376	rs176694	Zhernakova_WHOLEBLOOD-gene-primary	8.37E-25	G	-10.2834
ABO	1.74E-07	rs28645493	0.000417512	C	G	0.9031	-0.0328	rs28645493	Zhernakova_WHOLEBLOOD-gene-primary	8.86E-32	G	-11.7307
ABO	1.74E-07	rs4962050	0.00159826	C	G	0.3078	-0.0180	rs4962050	Zhernakova_WHOLEBLOOD-gene-primary	4.50E-25	C	10.343
ABO	1.74E-07	rs12683493	0.00756848	T	C	0.1990	-0.0166	rs12683493	Zhernakova_WHOLEBLOOD-gene-primary	1.39E-11	T	6.75868
ABO	1.74E-07	rs9411490	0.0892349	A	G	0.8214	0.0109	rs9411490	Zhernakova_WHOLEBLOOD-gene-contextspecific	2.57E-16	G	-8.19194
ABO	1.74E-07	rs12379977	0.265942	T	C	0.3299	-0.0061	rs12379977	Zhernakova_WHOLEBLOOD-gene-primary	2.72E-18	T	8.72254
ABO	1.74E-07	rs10901244	0.364805	T	C	0.4184	-0.0046	rs10901244	Hao_LUNG	0	T	-8.47
ABO	1.74E-07	rs626035	0.443795	T	G	0.8027	0.0048	rs626035	Zhernakova_WHOLEBLOOD-gene-primary	1.91E-17	G	8.49909
ABO	1.74E-07	rs9411463	0.492101	T	C	0.0867	-0.0069	rs9411463	Hao_LUNG	0	C	-11.7
ABO	1.74E-07	rs9411463	0.492101	T	C	0.0867	-0.0069	rs9411463	Zhernakova_WHOLEBLOOD-gene-primary	2.39E-73	T	18.1159
ABO	1.74E-07	rs9411463	0.492101	T	C	0.0867	-0.0069	rs9411463	GTE_SKIN	6.54E-20	T	1.12925
ABO	1.74E-07	rs9411463	0.492101	T	C	0.0867	-0.0069	rs9411463	GTE_SKIN	1.00E-17	T	1.38294
ABO	1.74E-07	rs9411463	0.492101	T	C	0.0867	-0.0069	rs9411463	GTE_LUNG	7.67E-12	T	0.626984
ABO	1.74E-07	rs35664240	0.592636	T	C	0.1105	0.0052	rs35664240	Zhernakova_WHOLEBLOOD-gene-primary	2.39E-12	T	-7.00975
ABO	1.74E-07	rs62576042	0.625485	T	C	0.1701	0.0037	rs62576042	Zhernakova_WHOLEBLOOD-gene-primary	2.18E-21	T	9.49601
ABO	1.74E-07	rs623361	0.635394	A	G	0.3350	0.0028	rs623361	Zhernakova_WHOLEBLOOD-gene-primary	4.77E-11	A	6.57812
ABO	1.74E-07	rs3094377	0.875632	T	C	0.0340	-0.0021	rs3094377	Zhernakova_WHOLEBLOOD-gene-primary	1.39E-18	T	-8.7985
AC004893.11	8.10E-07	rs4236540	8.10E-07	T	G	0.7143	-0.0282	rs4236540	Zhernakova_WHOLEBLOOD-exon-primary	3.67E-15	G	7.86572
AC004893.11	8.10E-07	rs4236540	8.10E-07	T	G	0.7143	-0.0282	rs4236540	Zhernakova_WHOLEBLOOD-gene-primary	1.16E-13	G	7.42145
AC004893.11	8.10E-07	rs4236540	8.10E-07	T	G	0.7143	-0.0282	rs4236540	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.16E-13	G	7.42145
APOBR	1.58E-08	rs151233	3.22E-06	T	C	0.1327	0.0385	rs151233	Zhernakova_WHOLEBLOOD-exon-primary	2.31E-12	T	-7.0144
APOBR	1.58E-08	rs151233	3.22E-06	T	C	0.1327	0.0385	rs151233	Zhernakova_WHOLEBLOOD-exon-primary	1.47E-11	T	-6.75084
APOBR	1.58E-08	rs151233	3.22E-06	T	C	0.1327	0.0385	rs151233	Zhernakova_WHOLEBLOOD-gene-contextspecific	9.17E-10	T	-6.12304
APOBR	1.58E-08	rs151233	3.22E-06	T	C	0.1327	0.0385	rs151233	Zhernakova_WHOLEBLOOD-gene-primary	9.17E-10	T	-6.12304
APOBR	1.58E-08	rs27741	4.42E-06	A	G	0.3622	-0.0239	rs27741	LloydJones_WHOLEBLOOD	9.90E-30	A	0.313874
ATXN2L	1.71E-06	rs8056890	1.71E-06	A	G	0.2959	-0.0254	rs8056890	Zhernakova_WHOLEBLOOD-gene-contextspecific	2.10E-23	A	-9.96813
ATXN2L	1.71E-06	rs8056890	1.71E-06	A	G	0.2959	-0.0254	rs8056890	Zhernakova_WHOLEBLOOD-gene-primary	2.10E-23	A	-9.96813
ATXN2L	1.71E-06	rs8056890	1.71E-06	A	G	0.2959	-0.0254	rs8056890	Zhernakova_WHOLEBLOOD-exon-primary	3.78E-22	A	-9.67673
ATXN2L	1.71E-06	rs8056890	1.71E-06	A	G	0.2959	-0.0254	rs8056890	Zhernakova_WHOLEBLOOD-gene-primary	1.27E-20	A	-9.31092
ATXN2L	1.71E-06	rs8056890	1.71E-06	A	G	0.2959	-0.0254	rs8056890	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.27E-20	A	-9.31092
ATXN2L	1.71E-06	rs8056890	1.71E-06	A	G	0.2959	-0.0254	rs8056890	Zhernakova_WHOLEBLOOD-exon-primary	4.79E-14	A	-7.53739
CASZ1	1.93E-06	rs12045923	4.87E-07	T	C	0.6497	0.0288	rs12045923	Kim_MONOCYTES-LPS	7.63E-12	NA	-9
CASZ1	1.93E-06	rs12045923	4.87E-07	T	C	0.6497	0.0288	rs12045923	Kim_MONOCYTES-Baseline	3.72E-11	NA	-9
CASZ1	1.93E-06	rs284296	0.29118	A	G	0.7823	-0.0070	rs284296	LloydJones_WHOLEBLOOD	7.10E-18	G	-0.282965
CASZ1	1.93E-06	rs284296	0.29118	A	G	0.7823	-0.0070	rs284296	Zhernakova_WHOLEBLOOD-exonratio-primary	3.41E-10	G	6.27892
CTD-2369P2.8	5.10E-07	rs1799969	1.76E-05	A	G	0.1480	0.0355	rs1799969	Zhernakova_WHOLEBLOOD-gene-primary	2.10E-14	A	-7.64456
CTD-2369P2.8	5.10E-07	rs78064630	0.00135441	A	G	0.0765	0.0314	rs78064630	Zhernakova_WHOLEBLOOD-gene-primary	2.46E-12	A	7.00583

Table E4

CTD-2369P2.8	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-exon-primary	2.96E-190	T	29.421
CTD-2369P2.8	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.17E-181	T	28.7411
CTD-2369P2.8	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-gene-primary	1.17E-181	T	28.7411
CTD-2369P2.8	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	GTE_WHOLEBLOOD	4.51E-21	T	0.410584
CTD-2369P2.8	5.10E-07	rs2633973	0.00276867	T	C	0.6599	-0.0168	rs2633973	Zhernakova_WHOLEBLOOD-gene-primary	8.83E-10	C	-6.12904
CTD-2369P2.8	5.10E-07	rs28378712	0.00678241	T	G	0.6803	-0.0152	rs28378712	Zhernakova_WHOLEBLOOD-gene-primary	3.55E-19	G	-8.94992
CTD-2369P2.8	5.10E-07	rs2017213	0.0075857	A	T	0.3810	-0.0139	rs2017213	Zhernakova_WHOLEBLOOD-gene-primary	6.90E-12	A	6.85981
CTD-2369P2.8	5.10E-07	rs10411056	0.0197799	A	C	0.9354	0.0226	rs10411056	Zhernakova_WHOLEBLOOD-gene-primary	2.83E-15	C	7.89849
CTD-2369P2.8	5.10E-07	rs79337061	0.348043	T	C	0.0765	-0.0091	rs79337061	Zhernakova_WHOLEBLOOD-gene-primary	3.58E-15	T	7.86884
CTD-2369P2.8	5.10E-07	rs2358581	0.352461	T	G	0.2483	-0.0056	rs2358581	Zhernakova_WHOLEBLOOD-gene-primary	1.63E-27	T	-10.8681
CTD-2369P2.8	5.10E-07	rs62130661	0.464875	C	G	0.8010	0.0047	rs62130661	Zhernakova_WHOLEBLOOD-gene-primary	6.30E-30	G	11.3644
CTD-2369P2.8	5.10E-07	rs77123125	0.758586	A	G	0.1003	-0.0032	rs77123125	Zhernakova_WHOLEBLOOD-gene-primary	7.17E-10	A	-6.16239
CTD-2369P2.8	5.10E-07	rs3177696	0.908246	T	C	0.9201	-0.0010	rs3177696	Zhernakova_WHOLEBLOOD-gene-primary	3.44E-15	C	7.874
FOSL2	8.76E-07	rs7562	8.76E-07	T	C	0.5306	0.0251	rs7562	Zhernakova_WHOLEBLOOD-exonratio-primary	2.60E-29	C	-11.2398
FOSL2	8.76E-07	rs7562	8.76E-07	T	C	0.5306	0.0251	rs7562	Zhernakova_WHOLEBLOOD-exonratio-primary	6.14E-26	C	10.5324
FOSL2	8.76E-07	rs7562	8.76E-07	T	C	0.5306	0.0251	rs7562	Walsh_WHOLEBLOOD	1.89E-11	T	-0.47246
FOSL2	8.76E-07	rs7562	8.76E-07	T	C	0.5306	0.0251	rs7562	Zhernakova_WHOLEBLOOD-exonratio-primary	1.41E-09	C	6.05473
FPR1	1.49E-08	rs7254019	1.79E-05	A	G	0.1395	0.0381	rs7254019	Zhernakova_WHOLEBLOOD-gene-primary	1.45E-10	A	-6.41041
FPR1	1.49E-08	rs12461801	9.65E-05	A	G	0.5969	-0.0203	rs12461801	Zhernakova_WHOLEBLOOD-exon-primary	1.83E-150	G	-26.1263
FPR1	1.49E-08	rs12461801	9.65E-05	A	G	0.5969	-0.0203	rs12461801	Zhernakova_WHOLEBLOOD-gene-primary	4.30E-132	G	-24.4563
FPR1	1.49E-08	rs12461801	9.65E-05	A	G	0.5969	-0.0203	rs12461801	Zhernakova_WHOLEBLOOD-exon-primary	6.10E-121	G	-23.3847
FPR1	1.49E-08	rs12461801	9.65E-05	A	G	0.5969	-0.0203	rs12461801	LloydJones_WHOLEBLOOD	7.10E-80	G	-0.532809
FPR1	1.49E-08	rs12461801	9.65E-05	A	G	0.5969	-0.0203	rs12461801	Zhernakova_WHOLEBLOOD-exon-primary	4.02E-45	G	-14.096
FPR1	1.49E-08	rs12461801	9.65E-05	A	G	0.5969	-0.0203	rs12461801	Zhernakova_WHOLEBLOOD-exon-primary	3.32E-20	G	-9.20782
FPR1	1.49E-08	rs12461801	9.65E-05	A	G	0.5969	-0.0203	rs12461801	Walsh_WHOLEBLOOD	5.91E-14	G	-0.538086
FPR1	1.49E-08	rs6509571	0.000191405	T	C	0.3520	-0.0198	rs6509571	Westra_WHOLEBLOOD	7.86E-73	T	18.0502
FPR1	1.49E-08	rs6509571	0.000191405	T	C	0.3520	-0.0198	rs6509571	LloydJones_WHOLEBLOOD	1.60E-38	T	0.373842
FPR1	1.49E-08	rs6509571	0.000191405	T	C	0.3520	-0.0198	rs6509571	Zhernakova_WHOLEBLOOD-gene-primary	3.51E-34	T	12.1901
FPR1	1.49E-08	rs7253284	0.000877773	A	G	0.3282	0.0180	rs7253284	Hao_LUNG	0	G	10.3
FPR1	1.49E-08	rs7253284	0.000877773	A	G	0.3282	0.0180	rs7253284	Zhernakova_WHOLEBLOOD-gene-primary	5.36E-93	A	-20.4556
FPR1	1.49E-08	rs7253284	0.000877773	A	G	0.3282	0.0180	rs7253284	Zeller_MONOCYTES	6.54E-92	NA	-9
FPR1	1.49E-08	rs7253284	0.000877773	A	G	0.3282	0.0180	rs7253284	Westra_WHOLEBLOOD	6.06E-90	A	-20.11
FPR1	1.49E-08	rs7253284	0.000877773	A	G	0.3282	0.0180	rs7253284	LloydJones_WHOLEBLOOD	3.50E-48	A	-0.441034
FPR1	1.49E-08	rs7253284	0.000877773	A	G	0.3282	0.0180	rs7253284	GTE_WHOLEBLOOD	1.35E-09	G	0.1661
FPR1	1.49E-08	rs62106945	0.00554085	T	C	0.0884	0.0252	rs62106945	Zhernakova_WHOLEBLOOD-gene-primary	1.73E-11	T	-6.7268
FPR1	1.49E-08	rs16983230	0.0164076	T	G	0.9558	-0.0316	rs16983230	Westra_WHOLEBLOOD	3.93E-10	G	-6.25665
FPR1	1.49E-08	rs79956121	0.0172967	T	C	0.0714	0.0277	rs79956121	Zhernakova_WHOLEBLOOD-gene-primary	2.37E-10	T	-6.33492
FPR1	1.49E-08	rs1868943	0.0584274	A	G	0.0986	0.0170	rs1868943	Zhernakova_WHOLEBLOOD-gene-primary	8.29E-31	A	-11.54
FPR1	1.49E-08	rs1868943	0.0584274	A	G	0.0986	0.0170	rs1868943	Zhernakova_WHOLEBLOOD-exon-primary	1.29E-19	A	-9.06152
FPR1	1.49E-08	rs17661285	0.0848097	A	G	0.3163	-0.0095	rs17661285	Westra_WHOLEBLOOD	3.20E-25	A	10.376
FPR1	1.49E-08	rs17661285	0.0848097	A	G	0.3163	-0.0095	rs17661285	LloydJones_WHOLEBLOOD	9.60E-13	A	0.211433
FPR1	1.49E-08	rs8110040	0.568977	A	G	0.0884	0.0050	rs8110040	Westra_WHOLEBLOOD	5.78E-17	A	-8.36968
FPR1	1.49E-08	rs8104640	0.573427	T	C	0.1207	0.0050	rs8104640	Westra_WHOLEBLOOD	4.18E-108	T	22.0876
FPR1	1.49E-08	rs8104640	0.573427	T	C	0.1207	0.0050	rs8104640	LloydJones_WHOLEBLOOD	1.00E-61	T	0.795929
FPR1	1.49E-08	rs8104640	0.573427	T	C	0.1207	0.0050	rs8104640	Walsh_WHOLEBLOOD	4.66E-28	C	1.2629
FPR1	1.49E-08	rs8104640	0.573427	T	C	0.1207	0.0050	rs8104640	GTE_FIBROBLASTS	7.55E-18	T	0.748262
FPR1	1.49E-08	rs17803207	0.624261	A	G	0.1122	-0.0037	rs17803207	Zhernakova_WHOLEBLOOD-exon-primary	2.36E-191	A	29.5069
FPR1	1.49E-08	rs17803207	0.624261	A	G	0.1122	-0.0037	rs17803207	Zhernakova_WHOLEBLOOD-exonratio-primary	2.23E-61	A	16.5302
FPR1	1.49E-08	rs17803207	0.624261	A	G	0.1122	-0.0037	rs17803207	Battle_WHOLEBLOOD-ase	4.72E-53	NA	-9
FPR1	1.49E-08	rs17803207	0.624261	A	G	0.1122	-0.0037	rs17803207	Zhernakova_WHOLEBLOOD-exon-primary	1.54E-12	A	7.07069
FPR1	1.49E-08	rs117223215	0.721224	T	C	0.0153	0.0055	rs117223215	LloydJones_WHOLEBLOOD	3.90E-20	T	-0.776936
FPR1	1.49E-08	rs73056872	0.919279	T	G	0.9133	-0.0009	rs73056872	LloydJones_WHOLEBLOOD	1.20E-15	G	-0.406202
ICAM1	5.10E-07	rs1799969	1.76E-05	A	G	0.1480	0.0355	rs1799969	Zhernakova_WHOLEBLOOD-gene-primary	2.10E-14	A	-7.64456
ICAM1	5.10E-07	rs78064630	0.00135441	A	G	0.0765	-0.0314	rs78064630	Zhernakova_WHOLEBLOOD-gene-primary	2.46E-12	A	7.00583
ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-exon-primary	2.96E-190	T	29.421
ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.17E-181	T	28.7411
ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-gene-primary	1.17E-181	T	28.7411

ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Raj_MONOCYTES	4.09E-94	NA	-9
ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Battle_WHOLEBLOOD	3.16E-44	NA	-9
ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-gene-primary	3.14E-30	T	11.4249
ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-exon-primary	1.62E-28	T	11.0773
ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-exon-primary	3.29E-26	T	10.5908
ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-exon-primary	7.79E-26	T	10.5098
ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-gene-primary	2.07E-21	T	9.50136
ICAM1	5.10E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-exon-primary	1.46E-13	T	7.39063
ICAM1	5.10E-07	rs2633973	0.00276867	T	C	0.6599	-0.0168	rs2633973	Zhernakova_WHOLEBLOOD-gene-primary	8.83E-10	C	-6.12904
ICAM1	5.10E-07	rs28378712	0.00678241	T	G	0.6803	-0.0152	rs28378712	Zhernakova_WHOLEBLOOD-gene-primary	3.55E-19	G	-8.94992
ICAM1	5.10E-07	rs2017213	0.0075857	A	T	0.3810	-0.0139	rs2017213	Zhernakova_WHOLEBLOOD-gene-primary	6.90E-12	A	6.85981
ICAM1	5.10E-07	rs10411056	0.0197799	A	C	0.9354	0.0226	rs10411056	Zhernakova_WHOLEBLOOD-gene-primary	2.83E-15	C	7.89849
ICAM1	5.10E-07	rs79337061	0.348043	T	C	0.0765	-0.0091	rs79337061	Zhernakova_WHOLEBLOOD-gene-primary	3.58E-15	T	7.86884
ICAM1	5.10E-07	rs2358581	0.352461	T	G	0.2483	-0.0056	rs2358581	Zhernakova_WHOLEBLOOD-gene-primary	1.63E-27	T	-10.8681
ICAM1	5.10E-07	rs62130661	0.464875	C	G	0.8010	0.0047	rs62130661	Zhernakova_WHOLEBLOOD-gene-primary	6.30E-30	G	11.3644
ICAM1	5.10E-07	rs77123125	0.758586	A	G	0.1003	-0.0032	rs77123125	Zhernakova_WHOLEBLOOD-gene-primary	7.17E-10	A	-6.16239
ICAM1	5.10E-07	rs3177696	0.908246	T	C	0.9201	-0.0010	rs3177696	Zhernakova_WHOLEBLOOD-gene-primary	3.44E-15	C	7.874
ICAM4	6.08E-07	rs1799969	1.76E-05	A	G	0.1480	0.0355	rs1799969	Zhernakova_WHOLEBLOOD-gene-primary	2.10E-14	A	-7.64456
ICAM4	6.08E-07	rs1799969	1.76E-05	A	G	0.1480	0.0355	rs1799969	Westra_WHOLEBLOOD	8.91E-11	A	-6.48453
ICAM4	6.08E-07	rs78064630	0.00135441	A	G	0.0765	-0.0314	rs78064630	Zhernakova_WHOLEBLOOD-gene-primary	2.46E-12	A	7.00583
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Battle_WHOLEBLOOD	1.06E-216	NA	-9
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-exon-primary	2.96E-190	T	29.421
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.17E-181	T	28.7411
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Zhernakova_WHOLEBLOOD-gene-primary	1.17E-181	T	28.7411
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Fairfax_MONOCYTES-NAIVE	2.35E-129	NA	-9
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Westra_WHOLEBLOOD	2.04E-120	T	23.3333
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	LloydJones_WHOLEBLOOD	9.10E-100	T	0.663692
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Fairfax_MONOCYTES-NAIVE	5.49E-97	NA	-9
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Fairfax_MONOCYTES	2.77E-92	NA	-9
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Fairfax_MONOCYTES	7.46E-62	NA	-9
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Westra_WHOLEBLOOD	8.86E-54	T	15.4397
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	LloydJones_WHOLEBLOOD	6.30E-38	T	0.400262
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Walsh_WHOLEBLOOD	1.28E-37	C	-0.967882
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	GTE_WHOLEBLOOD	1.83E-34	T	0.619615
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Grundberg_LCLS	1.43E-16	T	-0.100867
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Davenport_LEUCOCYTES	4.91E-11	NA	-9
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Yao_WHOLEBLOOD	2.19E-10	C	-0.045635
ICAM4	6.08E-07	rs281437	0.00184684	T	C	0.2670	-0.0178	rs281437	Yao_WHOLEBLOOD	2.19E-10	C	-0.045635
ICAM4	6.08E-07	rs2633973	0.00276867	T	C	0.6599	-0.0168	rs2633973	Zhernakova_WHOLEBLOOD-gene-primary	8.83E-10	C	-6.12904
ICAM4	6.08E-07	rs28378712	0.00678241	T	G	0.6803	-0.0152	rs28378712	Zhernakova_WHOLEBLOOD-gene-primary	3.55E-19	G	-8.94992
ICAM4	6.08E-07	rs2017213	0.0075857	A	T	0.3810	-0.0139	rs2017213	Zhernakova_WHOLEBLOOD-gene-primary	6.90E-12	A	6.85981
ICAM4	6.08E-07	rs10411056	0.0197799	A	C	0.9354	0.0226	rs10411056	Zhernakova_WHOLEBLOOD-gene-primary	2.83E-15	C	7.89849
ICAM4	6.08E-07	rs10411056	0.0197799	A	C	0.9354	0.0226	rs10411056	LloydJones_WHOLEBLOOD	9.30E-12	C	0.359728
ICAM4	6.08E-07	rs2358581	0.352461	T	G	0.2483	-0.0056	rs2358581	Zhernakova_WHOLEBLOOD-gene-primary	1.63E-27	T	-10.8681
ICAM4	6.08E-07	rs2358581	0.352461	T	G	0.2483	-0.0056	rs2358581	LloydJones_WHOLEBLOOD	1.00E-13	T	-0.222958
ICAM4	6.08E-07	rs7252450	0.397785	A	C	0.0782	-0.0082	rs7252450	LloydJones_WHOLEBLOOD	3.40E-15	A	0.413387
ICAM4	6.08E-07	rs7252450	0.397785	A	C	0.0782	-0.0082	rs7252450	Zhernakova_WHOLEBLOOD-gene-primary	1.19E-11	A	6.78206
ICAM4	6.08E-07	rs62130661	0.464875	C	G	0.8010	0.0047	rs62130661	Zhernakova_WHOLEBLOOD-gene-primary	6.30E-30	G	11.3644
ICAM4	6.08E-07	rs62130661	0.464875	C	G	0.8010	0.0047	rs62130661	LloydJones_WHOLEBLOOD	1.90E-16	G	0.285605
ICAM4	6.08E-07	rs77123125	0.758586	A	G	0.1003	-0.0032	rs77123125	Zhernakova_WHOLEBLOOD-gene-primary	7.17E-10	A	-6.16239
ICAM4	6.08E-07	rs3177696	0.908246	T	C	0.9201	-0.0010	rs3177696	Zhernakova_WHOLEBLOOD-gene-primary	3.44E-15	C	7.874
ICAM4	6.08E-07	rs3177696	0.908246	T	C	0.9201	-0.0010	rs3177696	LloydJones_WHOLEBLOOD	1.10E-10	C	0.314191
IL27	2.60E-08	rs231970	4.84E-06	A	G	0.8827	-0.0350	rs231970	Zhernakova_WHOLEBLOOD-gene-primary	4.80E-21	G	-9.41324
IL27	2.60E-08	rs7191548	1.18E-05	T	C	0.6599	0.0237	rs7191548	Zhernakova_WHOLEBLOOD-gene-primary	4.30E-36	C	12.5439
IL27	2.60E-08	rs7191548	1.18E-05	T	C	0.6599	0.0237	rs7191548	Zhernakova_WHOLEBLOOD-gene-contextspecific	4.30E-36	C	12.5439
IL27	2.60E-08	rs7191548	1.18E-05	T	C	0.6599	0.0237	rs7191548	Zhernakova_WHOLEBLOOD-exon-primary	3.14E-20	C	9.21402
IPCEF1	4.33E-07	rs9397706	4.73E-07	A	G	0.5272	0.0257	rs9397706	Zhernakova_WHOLEBLOOD-gene-primary	2.74E-13	G	7.30675

IPCEF1	4.33E-07	rs7759388	0.00300347	A	G	0.1599	0.0200	rs7759388	Zhernakova_WHOLEBLOOD-exon-primary	1.51E-23	A	-10.0008
IPCEF1	4.33E-07	rs7759388	0.00300347	A	G	0.1599	0.0200	rs7759388	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.99E-19	A	-9.01372
IPCEF1	4.33E-07	rs7759388	0.00300347	A	G	0.1599	0.0200	rs7759388	Zhernakova_WHOLEBLOOD-gene-primary	1.99E-19	A	-9.01372
IPCEF1	4.33E-07	rs7759388	0.00300347	A	G	0.1599	0.0200	rs7759388	Battle_WHOLEBLOOD	2.78E-10	NA	-9
IPCEF1	4.33E-07	rs1406055	0.0498256	C	G	0.3214	0.0110	rs1406055	Zhernakova_WHOLEBLOOD-exon-primary	2.26E-153	C	26.3811
IPCEF1	4.33E-07	rs1406055	0.0498256	C	G	0.3214	0.0110	rs1406055	Zhernakova_WHOLEBLOOD-exonratio-primary	5.96E-128	C	24.0642
IPCEF1	4.33E-07	rs12528985	0.676521	T	C	0.3486	-0.0021	rs12528985	GTE_SKIN	1.39E-10	T	-0.363454
LAT	1.06E-06	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Fehrmann_WHOLEBLOOD	5.40E-177	G	28.37
LAT	1.06E-06	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-gene-primary	4.73E-19	C	8.91852
LAT	1.06E-06	rs6565261	0.000208433	A	C	0.2721	0.0212	rs6565261	Fehrmann_WHOLEBLOOD	2.40E-21	A	-9.48
LAT	1.06E-06	rs4788115	0.122264	A	T	0.2160	0.0112	rs4788115	Zhernakova_WHOLEBLOOD-exon-primary	3.58E-23	A	9.91507
LAT	1.06E-06	rs4788115	0.122264	A	T	0.2160	0.0112	rs4788115	Zhernakova_WHOLEBLOOD-exon-primary	1.57E-16	A	-8.25109
LAT	1.06E-06	rs4788115	0.122264	A	T	0.2160	0.0112	rs4788115	Zhernakova_WHOLEBLOOD-gene-primary	3.93E-10	A	-6.25655
MYO1C	1.52E-06	rs56157500	9.23E-05	A	C	0.7738	-0.0239	rs56157500	Zhernakova_WHOLEBLOOD-gene-primary	3.06E-12	C	6.97484
MYO1C	1.52E-06	rs59343110	9.79E-05	T	C	0.0748	-0.0338	rs59343110	Zhernakova_WHOLEBLOOD-gene-primary	1.82E-13	T	7.36164
MYO1C	1.52E-06	rs59343110	9.79E-05	T	C	0.0748	-0.0338	rs59343110	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.84E-11	T	6.71784
MYO1C	1.52E-06	rs147003567	0.0395452	A	G	0.0595	0.0206	rs147003567	Zhernakova_WHOLEBLOOD-gene-contextspecific	2.26E-22	A	-9.72914
MYO1C	1.52E-06	rs147003567	0.0395452	A	G	0.0595	0.0206	rs147003567	Zhernakova_WHOLEBLOOD-gene-primary	2.26E-22	A	-9.72914
MYO1C	1.52E-06	rs147003567	0.0395452	A	G	0.0595	0.0206	rs147003567	Zhernakova_WHOLEBLOOD-exon-primary	1.42E-20	A	-9.29855
MYO1C	1.52E-06	rs4790152	0.386692	T	C	0.5527	-0.0045	rs4790152	Hao_LUNG	0	G	-8.3
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Fairfax_MONOCYTES-LPS24	6.05E-40	NA	-9
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Fairfax_MONOCYTES-LPS24	7.82E-35	NA	-9
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zeller_MONOCYTES	2.25E-28	NA	-9
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Westra_WHOLEBLOOD	3.24E-28	C	-11.015
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Fairfax_MONOCYTES-LPS24	8.02E-15	NA	-9
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Westra_WHOLEBLOOD	6.52E-13	C	-7.18937
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Raj_MONOCYTES	1.76E-12	NA	-9
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Raj_CD4TCELLS	4.44E-12	NA	-9
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	LloydJones_WHOLEBLOOD	1.20E-11	C	-0.199056
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Quach_MONOCYTES-baseline	1.93E-11	NA	0.4
NCF4	4.85E-08	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zhernakova_WHOLEBLOOD-gene-primary	1.85E-09	C	-6.01071
NCF4	4.85E-08	rs5756391	8.53E-06	A	G	0.3520	-0.0232	rs5756391	Westra_WHOLEBLOOD	3.81E-16	A	8.14455
NCF4	4.85E-08	rs4821549	0.0746725	A	G	0.0646	-0.0182	rs4821549	Raj_CD4TCELLS	8.38E-29	NA	-9
NCF4	4.85E-08	rs2072710	0.329012	A	G	0.2874	-0.0057	rs2072710	Raj_CD4TCELLS	3.43E-70	NA	-9
NCF4	4.85E-08	rs2072710	0.329012	A	G	0.2874	-0.0057	rs2072710	Westra_WHOLEBLOOD	3.55E-36	A	12.5591
NCF4	4.85E-08	rs2072710	0.329012	A	G	0.2874	-0.0057	rs2072710	Kasela_CD4TCELLS	1.34E-19	A	9.05737
NCF4	4.85E-08	rs2072710	0.329012	A	G	0.2874	-0.0057	rs2072710	Fairfax_MONOCYTES-LPS24	2.13E-13	NA	-9
NCF4	4.85E-08	rs2072710	0.329012	A	G	0.2874	-0.0057	rs2072710	Fairfax_MONOCYTES-LPS24	2.29E-11	NA	-9
NCF4	4.85E-08	rs2072710	0.329012	A	G	0.2874	-0.0057	rs2072710	Westra_WHOLEBLOOD	1.11E-09	A	-6.0922
NCF4	4.85E-08	rs5756363	0.353738	T	G	0.1276	-0.0067	rs5756363	Westra_WHOLEBLOOD	3.61E-10	T	-6.27023
NSMCE1	6.62E-07	rs4523932	5.15E-06	A	G	0.0119	-0.1047	rs4523932	Nedelec_MACROPHAGES-listeria-asQTL	1.40E-22	NA	0.130657
NSMCE1	6.62E-07	rs4523932	5.15E-06	A	G	0.0119	-0.1047	rs4523932	Nedelec_MACROPHAGES-baseline-asQTL	1.04E-17	NA	-0.11553
NSMCE1	6.62E-07	rs9788909	0.00370886	A	G	0.3452	0.0157	rs9788909	Zeller_MONOCYTES	1.72E-28	NA	-9
NSMCE1	6.62E-07	rs9788909	0.00370886	A	G	0.3452	0.0157	rs9788909	LloydJones_WHOLEBLOOD	2.00E-12	A	0.206931
NUP43	2.08E-06	rs6909158	7.46E-06	T	C	0.8078	-0.0297	rs9478311	Westra_WHOLEBLOOD	8.60E-11	G	-6.48984
NUP43	2.08E-06	rs4870139	1.29E-05	T	C	0.6412	-0.0236	rs4870139	Hao_LUNG	0	T	8.36
NUP43	2.08E-06	rs4870139	1.29E-05	T	C	0.6412	-0.0236	rs4870139	LloydJones_WHOLEBLOOD	1.70E-149	C	-0.738124
NUP43	2.08E-06	rs4870139	1.29E-05	T	C	0.6412	-0.0236	rs4870139	Zhernakova_WHOLEBLOOD-gene-primary	1.88E-97	C	-20.9498
NUP43	2.08E-06	rs4870139	1.29E-05	T	C	0.6412	-0.0236	rs4870139	Fairfax_MONOCYTES-NAIVE	4.66E-74	NA	-9
NUP43	2.08E-06	rs4870139	1.29E-05	T	C	0.6412	-0.0236	rs4870139	Raj_CD4TCELLS	4.00E-53	NA	-9
NUP43	2.08E-06	rs4870139	1.29E-05	T	C	0.6412	-0.0236	rs4870139	GTE_WHOLEBLOOD	8.91E-20	C	-0.422135
NUP43	2.08E-06	rs4870139	1.29E-05	T	C	0.6412	-0.0236	rs4870139	LloydJones_WHOLEBLOOD	3.10E-15	C	-0.223493
NUP43	2.08E-06	rs4870139	1.29E-05	T	C	0.6412	-0.0236	rs4870139	Davenport_LEUCOCYTES	2.18E-14	NA	-9
NUP43	2.08E-06	rs4870139	1.29E-05	T	C	0.6412	-0.0236	rs4870139	Kim_MONOCYTES-LPS	9.04E-11	NA	-9
NUP43	2.08E-06	rs4870139	1.29E-05	T	C	0.6412	-0.0236	rs4870139	Kasela_CD8TCELLS	1.89E-10	C	-6.37023
NUP43	2.08E-06	rs237004	0.000352558	A	C	0.3316	0.0190	rs237004	Fairfax_MONOCYTES-LPS24	1.01E-11	NA	-9
NUP43	2.08E-06	rs62439806	0.00671963	A	G	0.0408	0.0371	rs62439806	Raj_CD4TCELLS	6.89E-32	NA	-9

Table E4

NUP43	2.08E-06	rs62439806	0.00671963	A	G	0.0408	0.0371	rs62439806	LloydJones_WHOLEBLOOD	1.90E-10	A	-0.545214
NUP43	2.08E-06	rs12523685	0.0240397	A	T	0.7942	-0.0136	rs12523685	LloydJones_WHOLEBLOOD	1.20E-09	T	-0.191625
NUP43	2.08E-06	rs2281436	0.0272181	T	C	0.2874	0.0128	rs2281436	LloydJones_WHOLEBLOOD	3.60E-10	T	-0.191296
NUP43	2.08E-06	rs9379347	0.102521	A	G	0.0204	-0.0297	rs9379347	Raj_CD4TCELLS	8.84E-10	NA	-9
NUP43	2.08E-06	rs78765468	0.117714	C	G	0.0969	-0.0152	rs78765468	LloydJones_WHOLEBLOOD	9.20E-14	C	0.350226
NUP43	2.08E-06	rs112165893	0.16007	C	G	0.9813	-0.0288	rs113263746	Raj_CD4TCELLS	2.49E-32	NA	-9
NUP43	2.08E-06	rs112165893	0.16007	C	G	0.9813	-0.0288	rs113263746	Raj_MONOCYTES	3.67E-26	NA	-9
NUP43	2.08E-06	rs117067995	0.161872	A	G	0.0510	-0.0181	rs117067995	Raj_MONOCYTES	3.61E-18	NA	-9
NUP43	2.08E-06	rs117067995	0.161872	A	G	0.0510	-0.0181	rs117067995	Raj_CD4TCELLS	1.95E-11	NA	-9
NUP43	2.08E-06	rs13205080	0.186375	T	C	0.0544	-0.0152	rs13205080	Raj_MONOCYTES	6.05E-13	NA	-9
NUP43	2.08E-06	rs13205080	0.186375	T	C	0.0544	-0.0152	rs13205080	Raj_CD4TCELLS	1.67E-10	NA	-9
NUP43	2.08E-06	rs2151910	0.2168	A	T	0.2500	-0.0074	rs2151910	Westra_WHOLEBLOOD	1.09E-41	A	13.5268
NUP43	2.08E-06	rs2151910	0.2168	A	T	0.2500	-0.0074	rs2151910	LloydJones_WHOLEBLOOD	3.50E-28	A	0.341155
NUP43	2.08E-06	rs2151910	0.2168	A	T	0.2500	-0.0074	rs2151910	Zhernakova_WHOLEBLOOD-gene-primary	9.44E-17	A	8.31178
NUP43	2.08E-06	rs7762285	0.266123	A	G	0.8622	0.0092	rs7762285	Westra_WHOLEBLOOD	3.41E-15	G	7.87491
NUP43	2.08E-06	rs4870058	0.315372	A	G	0.0544	-0.0145	rs4870058	Raj_CD4TCELLS	2.16E-29	NA	-9
NUP43	2.08E-06	rs4870058	0.315372	A	G	0.0544	-0.0145	rs4870058	LloydJones_WHOLEBLOOD	7.10E-16	A	-0.534455
NUP43	2.08E-06	rs17733403	0.358474	A	G	0.0646	-0.0093	rs17733403	Westra_WHOLEBLOOD	3.24E-10	A	6.28695
NUP43	2.08E-06	rs6922028	0.498364	T	G	0.4626	0.0035	rs6922028	LloydJones_WHOLEBLOOD	3.20E-10	T	0.169703
NUP43	2.08E-06	rs60328093	0.546268	T	C	0.0102	-0.0200	rs200592260	GTE_WHOLEBLOOD	7.31E-19	A	-0.39905
NUP43	2.08E-06	rs60328093	0.546268	T	C	0.0102	-0.0200	rs200592260	GTE_SKIN	2.75E-10	A	-0.323661
NUP43	2.08E-06	rs9942443	0.590615	A	C	0.2840	0.0031	rs9942443	LloydJones_WHOLEBLOOD	5.50E-13	A	-0.222816
NUP43	2.08E-06	rs139394852	0.652985	A	C	0.9235	-0.0051	rs139394852	Raj_CD4TCELLS	1.25E-16	NA	-9
NUP43	2.08E-06	rs139394852	0.652985	A	C	0.9235	-0.0051	rs139394852	Raj_MONOCYTES	5.90E-14	NA	-9
NUP43	2.08E-06	rs2789503	0.681245	A	G	0.4949	-0.0021	rs2789503	Westra_WHOLEBLOOD	3.28E-30	G	-11.4212
NUP43	2.08E-06	rs2789503	0.681245	A	G	0.4949	-0.0021	rs2789503	LloydJones_WHOLEBLOOD	5.60E-15	G	-0.211463
NUP43	2.08E-06	rs2789503	0.681245	A	G	0.4949	-0.0021	rs2789503	Fairfax_MONOCYTES-NAIVE	1.88E-10	NA	-9
NUP43	2.08E-06	rs79782857	0.727052	A	T	0.9864	-0.0070	rs79782857	Raj_CD4TCELLS	4.35E-13	NA	-9
NUP43	2.08E-06	rs6918774	0.935752	T	C	0.1327	-0.0007	rs6918774	LloydJones_WHOLEBLOOD	1.50E-11	T	-0.288587
OR10J5	2.26E-06	rs2427837	1.02E-05	A	G	0.2942	-0.0248	rs2427837	Westra_WHOLEBLOOD	6.38E-24	A	10.0858
OR10J5	2.26E-06	rs2494260	0.00142189	T	C	0.2211	0.0186	rs2494260	Westra_WHOLEBLOOD	4.32E-13	T	-7.24505
OR10J5	2.26E-06	rs6699459	0.00911848	A	G	0.2058	-0.0160	rs6699459	Westra_WHOLEBLOOD	8.31E-17	A	8.32692
PRR5L	9.72E-07	rs7925585	2.78E-06	A	G	0.6565	0.0244	rs7925585	Zhernakova_WHOLEBLOOD-gene-contextspecific	3.23E-12	G	6.96757
PRR5L	9.72E-07	rs7925585	2.78E-06	A	G	0.6565	0.0244	rs7925585	Zhernakova_WHOLEBLOOD-exon-primary	2.35E-11	G	6.68226
PRR5L	9.72E-07	rs7925585	2.78E-06	A	G	0.6565	0.0244	rs7925585	Zhernakova_WHOLEBLOOD-gene-primary	3.72E-11	G	6.6152
PRR5L	9.72E-07	rs12270539	0.000107622	T	C	0.0374	0.0593	rs12270539	LloydJones_WHOLEBLOOD	2.90E-27	T	-0.819609
PRR5L	9.72E-07	rs429034	0.117757	A	T	0.7721	0.0099	rs429034	Zhernakova_WHOLEBLOOD-gene-primary	4.15E-19	T	-8.93263
PRR5L	9.72E-07	rs429034	0.117757	A	T	0.7721	0.0099	rs429034	Zhernakova_WHOLEBLOOD-exon-primary	4.02E-17	T	-8.41232
PRR5L	9.72E-07	rs429034	0.117757	A	T	0.7721	0.0099	rs429034	Zhernakova_WHOLEBLOOD-gene-contextspecific	3.41E-16	T	-8.15788
PRR5L	9.72E-07	rs1123347	0.367446	T	G	0.4303	0.0046	rs1123347	Zhernakova_WHOLEBLOOD-exonratio-primary	1.96E-59	T	16.2581
PRR5L	9.72E-07	rs1123347	0.367446	T	G	0.4303	0.0046	rs1123347	Zhernakova_WHOLEBLOOD-exonratio-primary	3.01E-35	T	-12.3886
PRR5L	9.72E-07	rs1123347	0.367446	T	G	0.4303	0.0046	rs1123347	Zhernakova_WHOLEBLOOD-exonratio-primary	3.03E-34	T	-12.2021
PRR5L	9.72E-07	rs1123347	0.367446	T	G	0.4303	0.0046	rs1123347	Zhernakova_WHOLEBLOOD-exonratio-primary	1.99E-25	T	-10.421
PRR5L	9.72E-07	rs1123347	0.367446	T	G	0.4303	0.0046	rs1123347	Zhernakova_WHOLEBLOOD-gene-primary	2.37E-21	T	9.4876
PRR5L	9.72E-07	rs1123347	0.367446	T	G	0.4303	0.0046	rs1123347	Zhernakova_WHOLEBLOOD-exonratio-primary	3.77E-19	T	-8.94371
PRR5L	9.72E-07	rs1123347	0.367446	T	G	0.4303	0.0046	rs1123347	Zhernakova_WHOLEBLOOD-exon-primary	1.48E-14	T	7.68951
PRR5L	9.72E-07	rs1123347	0.367446	T	G	0.4303	0.0046	rs1123347	Zhernakova_WHOLEBLOOD-exonratio-primary	1.88E-13	T	-7.35723
PRR5L	9.72E-07	rs1123347	0.367446	T	G	0.4303	0.0046	rs1123347	Zhernakova_WHOLEBLOOD-exonratio-primary	5.43E-11	T	-6.55878
PRR5L	9.72E-07	rs12785381	0.492162	T	C	0.0986	0.0061	rs12785381	Zhernakova_WHOLEBLOOD-exonratio-primary	1.06E-28	T	-11.1153
PRR5L	9.72E-07	rs12785381	0.492162	T	C	0.0986	0.0061	rs12785381	Zhernakova_WHOLEBLOOD-exon-primary	1.95E-27	T	-10.8518
PRR5L	9.72E-07	rs10836545	0.82493	A	G	0.9082	0.0021	rs10836545	Zhernakova_WHOLEBLOOD-exon-primary	5.84E-10	G	6.19496
PTPLA	8.16E-07	rs7092926	6.75E-07	T	G	0.4320	0.0259	rs7092926	Zhernakova_WHOLEBLOOD-gene-primary	4.91E-12	T	-6.90808
PTPLA	8.16E-07	rs7092926	6.75E-07	T	G	0.4320	0.0259	rs7092926	Zhernakova_WHOLEBLOOD-gene-contextspecific	4.91E-12	T	-6.90808
PTPLA	8.16E-07	rs17141430	0.0127591	A	G	0.8759	-0.0203	rs17141430	Westra_WHOLEBLOOD	3.55E-10	G	-6.27279
PTPLA	8.16E-07	rs78704091	0.0276797	A	T	0.0680	-0.0184	rs78704091	Zhernakova_WHOLEBLOOD-gene-primary	1.05E-11	A	6.79888
PTPLA	8.16E-07	rs7094705	0.0361169	A	G	0.3844	0.0111	rs7094705	Westra_WHOLEBLOOD	6.95E-23	A	9.84877
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	LloydJones_WHOLEBLOOD	1.20E-243	C	-0.971171

Table E4

PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.61E-202	C	-30.3645
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zhernakova_WHOLEBLOOD-gene-primary	1.61E-202	C	-30.3645
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zhernakova_WHOLEBLOOD-exon-primary	9.85E-174	C	-28.0997
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zhernakova_WHOLEBLOOD-exon-primary	1.58E-168	C	-27.6705
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zhernakova_WHOLEBLOOD-exon-primary	3.02E-154	C	-26.457
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zeller_MONOCYTES	9.09E-150	NA	-9
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Battle_WHOLEBLOOD	7.06E-135	NA	-9
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zhernakova_WHOLEBLOOD-exon-primary	6.01E-126	C	-23.8717
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Fehrmann_WHOLEBLOOD	8.30E-100	G	-21.21
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Jansen_WHOLEBLOOD	4.40E-93	C	-0.436
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Walsh_WHOLEBLOOD	1.45E-28	T	0.810608
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	GTE_WHOLEBLOOD	2.20E-27	C	-0.575968
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Battle_WHOLEBLOOD-ase	1.58E-21	NA	-9
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zhernakova_WHOLEBLOOD-exonratio-primary	3.09E-18	C	-8.70802
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Fehrmann_WHOLEBLOOD	2.10E-17	C	-8.49
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Davenport_LEUCOCYTES	4.57E-15	NA	-9
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Dinarzo_WHOLEBLOOD	1.07E-11	T	0.993866
PVALB	1.27E-07	rs4821544	3.58E-07	T	C	0.6854	-0.0291	rs4821544	Zhernakova_WHOLEBLOOD-polyAratio-primary	2.07E-11	C	-6.70078
PVALB	1.27E-07	rs1015775	0.000118906	A	G	0.1769	-0.0263	rs1015775	LloydJones_WHOLEBLOOD	8.20E-69	A	0.625638
PVALB	1.27E-07	rs1015775	0.000118906	A	G	0.1769	-0.0263	rs1015775	Zhernakova_WHOLEBLOOD-gene-primary	2.79E-44	A	13.9587
PVALB	1.27E-07	rs1015775	0.000118906	A	G	0.1769	-0.0263	rs1015775	Walsh_WHOLEBLOOD	3.33E-14	A	0.711935
PVALB	1.27E-07	rs730483	0.0661628	A	C	0.0816	0.0163	rs730483	LloydJones_WHOLEBLOOD	5.00E-19	A	-0.397495
PVALB	1.27E-07	rs730483	0.0661628	A	C	0.0816	0.0163	rs730483	Zhernakova_WHOLEBLOOD-gene-primary	1.72E-18	A	-8.77384
PVALB	1.27E-07	rs74971025	0.0668636	A	G	0.8435	-0.0135	rs74971025	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.37E-12	G	-7.08692
PVALB	1.27E-07	rs74971025	0.0668636	A	G	0.8435	-0.0135	rs74971025	LloydJones_WHOLEBLOOD	4.60E-11	G	-0.258926
PVALB	1.27E-07	rs74971025	0.0668636	A	G	0.8435	-0.0135	rs74971025	Zhernakova_WHOLEBLOOD-exon-primary	9.22E-11	G	-6.47915
PVALB	1.27E-07	rs74971025	0.0668636	A	G	0.8435	-0.0135	rs74971025	Zhernakova_WHOLEBLOOD-gene-primary	4.15E-10	G	-6.24834
PVALB	1.27E-07	rs4820250	0.342696	A	G	0.1344	-0.0068	rs4820250	Fairfax_MONOCYTES-LPS24	9.04E-15	NA	-9
RBM15B	1.41E-07	rs73078636	1.41E-07	A	G	0.1480	-0.0403	rs73078636	Zhernakova_WHOLEBLOOD-exon-primary	2.14E-52	A	15.233
RBM15B	1.41E-07	rs73078636	1.41E-07	A	G	0.1480	-0.0403	rs73078636	Zhernakova_WHOLEBLOOD-gene-contextspecific	2.08E-44	A	13.9794
RBM15B	1.41E-07	rs73078636	1.41E-07	A	G	0.1480	-0.0403	rs73078636	Zhernakova_WHOLEBLOOD-gene-primary	2.08E-44	A	13.9794
RBM15B	1.41E-07	rs73078636	1.41E-07	A	G	0.1480	-0.0403	rs73078636	Jansen_WHOLEBLOOD	8.50E-13	A	0.146
RP11-24N18.1	1.71E-06	rs8056890	1.71E-06	A	G	0.2959	-0.0254	rs8056890	Zhernakova_WHOLEBLOOD-exon-primary	3.78E-22	A	-9.67673
RP11-24N18.1	1.71E-06	rs8056890	1.71E-06	A	G	0.2959	-0.0254	rs8056890	Zhernakova_WHOLEBLOOD-gene-primary	1.27E-20	A	-9.31092
RP11-24N18.1	1.71E-06	rs8056890	1.71E-06	A	G	0.2959	-0.0254	rs8056890	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.27E-20	A	-9.31092
RP11-264B17.4	1.05E-06	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-exon-primary	3.51E-130	C	24.276
RP11-264B17.4	1.05E-06	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-gene-primary	4.38E-116	C	22.9027
RP11-264B17.4	1.05E-06	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-gene-contextspecific	4.38E-116	C	22.9027
RP11-264B17.4	1.05E-06	rs67479058	0.000161762	T	C	0.8503	-0.0258	rs67479058	Zhernakova_WHOLEBLOOD-gene-primary	2.55E-13	C	-7.3162
RP11-264B17.4	1.05E-06	rs12448482	0.197731	A	G	0.3861	0.0071	rs12448482	Zhernakova_WHOLEBLOOD-gene-primary	2.51E-24	A	-10.1769
RP11-534L20.5	1.32E-06	rs11117858	8.09E-06	T	G	0.7840	-0.0287	rs11117858	Zhernakova_WHOLEBLOOD-gene-primary	4.88E-54	G	-15.4778
RP11-534L20.5	1.32E-06	rs11117858	8.09E-06	T	G	0.7840	-0.0287	rs11117858	Kasela_CD4TCELLS	7.94E-11	G	-6.5018
RP11-534L20.5	1.32E-06	rs11117858	8.09E-06	T	G	0.7840	-0.0287	rs11117858	Kasela_CD8TCELLS	1.12E-09	G	-6.09141
RP11-534L20.5	1.32E-06	rs874718	5.06E-05	A	C	0.4728	0.0211	rs874718	Zhernakova_WHOLEBLOOD-gene-primary	5.77E-288	C	36.2652
RP11-534L20.5	1.32E-06	rs874718	5.06E-05	A	C	0.4728	0.0211	rs874718	Walsh_WHOLEBLOOD	1.69E-62	A	-0.975764
RP11-534L20.5	1.32E-06	rs874718	5.06E-05	A	C	0.4728	0.0211	rs874718	Kasela_CD4TCELLS	3.09E-49	A	-14.7498
RP11-534L20.5	1.32E-06	rs874718	5.06E-05	A	C	0.4728	0.0211	rs874718	Kasela_CD8TCELLS	2.55E-43	A	-13.8
RP11-534L20.5	1.32E-06	rs874718	5.06E-05	A	C	0.4728	0.0211	rs874718	GTE_SKIN	1.09E-19	C	0.690688
RP11-534L20.5	1.32E-06	rs874718	5.06E-05	A	C	0.4728	0.0211	rs874718	GTE_WHOLEBLOOD	1.57E-17	C	0.600601
RP11-534L20.5	1.32E-06	rs874718	5.06E-05	A	C	0.4728	0.0211	rs874718	GTE_SKIN	7.58E-16	C	0.743489
RP11-534L20.5	1.32E-06	rs874718	5.06E-05	A	C	0.4728	0.0211	rs874718	GTE_FIBROBLASTS	1.21E-11	C	0.564444
RP11-534L20.5	1.32E-06	rs874718	5.06E-05	A	C	0.4728	0.0211	rs874718	GTE_LUNG	7.63E-10	C	0.545556
RP11-534L20.5	1.32E-06	rs2987936	0.000433746	C	G	0.5357	-0.0176	rs2987936	Zhernakova_WHOLEBLOOD-gene-primary	5.63E-70	G	-17.6833
RP11-534L20.5	1.32E-06	rs2987936	0.000433746	C	G	0.5357	-0.0176	rs2987936	Kasela_CD4TCELLS	5.97E-14	G	-7.5086
RP11-534L20.5	1.32E-06	rs2987936	0.000433746	C	G	0.5357	-0.0176	rs2987936	Kasela_CD8TCELLS	6.02E-12	G	-6.87903
RP11-534L20.5	1.32E-06	rs2297546	0.00571061	C	G	0.5901	-0.0158	rs2297546	Zhernakova_WHOLEBLOOD-gene-primary	1.52E-79	G	-18.8846
RP11-534L20.5	1.32E-06	rs41299005	0.0199849	T	C	0.0663	-0.0209	rs41299005	Zhernakova_WHOLEBLOOD-gene-primary	2.23E-39	T	13.1298

Table E4

RP11-534L20.5	1.32E-06	rs35785716	0.060621	T	C	0.2517	-0.0113	rs35785716	Zhernakova_WHOLEBLOOD-gene-primary	6.88E-28	T	10.9469
RP11-534L20.5	1.32E-06	rs35785716	0.060621	T	C	0.2517	-0.0113	rs35785716	Kasela_CD8TCELLS	2.13E-10	T	6.35156
RP11-534L20.5	1.32E-06	rs2336941	0.0952277	T	C	0.1684	0.0109	rs2336941	Zhernakova_WHOLEBLOOD-gene-primary	2.35E-107	T	-22.0092
RP11-534L20.5	1.32E-06	rs2336941	0.0952277	T	C	0.1684	0.0109	rs2336941	Kasela_CD4TCELLS	1.45E-24	T	-10.2303
RP11-534L20.5	1.32E-06	rs2336941	0.0952277	T	C	0.1684	0.0109	rs2336941	Kasela_CD8TCELLS	2.63E-19	T	-8.98323
RP11-534L20.5	1.32E-06	rs2336941	0.0952277	T	C	0.1684	0.0109	rs2336941	Walsh_WHOLEBLOOD	1.23E-12	T	-0.621843
RP11-534L20.5	1.32E-06	rs7538261	0.112295	A	G	0.0629	0.0201	rs7538261	Zhernakova_WHOLEBLOOD-gene-primary	8.71E-17	A	-8.32111
RP11-534L20.5	1.32E-06	rs74882519	0.374648	A	G	0.0782	-0.0116	rs74882519	Zhernakova_WHOLEBLOOD-gene-primary	6.63E-51	A	15.0069
RP11-534L20.5	1.32E-06	rs6666087	0.376914	T	C	0.5748	-0.0046	rs6666087	Zhernakova_WHOLEBLOOD-gene-primary	2.96E-12	C	-6.97923
RP11-534L20.5	1.32E-06	rs61816859	0.415005	T	C	0.0578	0.0109	rs61816859	Zhernakova_WHOLEBLOOD-gene-primary	1.36E-14	T	-7.69955
RP11-534L20.5	1.32E-06	rs17433909	0.70506	T	C	0.0612	0.0051	rs17433909	Zhernakova_WHOLEBLOOD-gene-primary	1.11E-12	T	-7.11548
SPNS1	3.51E-09	rs2726040	4.84E-06	C	G	0.5748	-0.0238	rs2726040	LloydJones_WHOLEBLOOD	5.10E-40	G	-0.374583
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-gene-primary	7.23E-258	C	34.3026
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-gene-contextspecific	7.23E-258	C	34.3026
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-exon-primary	1.54E-237	C	32.9109
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	LloydJones_WHOLEBLOOD	9.20E-233	C	0.959846
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Battle_WHOLEBLOOD	9.62E-226	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Westra_WHOLEBLOOD	9.81E-198	C	50.796
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Fehrmann_WHOLEBLOOD	5.40E-177	G	28.37
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-exon-primary	8.15E-166	C	27.4443
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-exon-primary	6.31E-153	C	26.3424
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-exon-primary	7.74E-143	C	25.4465
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zeller_MONOCYTES	3.69E-133	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-exon-primary	3.51E-130	C	24.276
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-gene-primary	4.38E-116	C	22.9027
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-gene-contextspecific	4.38E-116	C	22.9027
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-exon-primary	5.38E-107	C	21.9716
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-exon-primary	3.90E-94	C	20.5831
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Fairfax_MONOCYTES-LPS24	1.23E-73	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Davenport_LEUCOCYTES	5.62E-30	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Grundberg_LCLS	2.72E-29	T	-0.153062
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	GTE_WHOLEBLOOD	1.93E-24	T	-0.284129
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Kim_MONOCYTES-LPS	2.25E-22	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Raj_MONOCYTES	2.62E-22	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	1.37E-21	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	1.91E-21	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	5.01E-21	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	1.50E-20	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	2.07E-20	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	1.41E-19	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Zhernakova_WHOLEBLOOD-gene-primary	4.73E-19	C	8.91852
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Kim_MONOCYTES-Baseline	8.22E-18	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	1.10E-16	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	6.93E-16	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	1.50E-14	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	GTE_LCLS	8.03E-14	T	-0.841903
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	1.04E-13	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Geuvadis_LCLS	4.34E-13	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Raj_CD4TCELLS	3.76E-10	NA	-9
SPNS1	3.51E-09	rs8045689	6.27E-06	T	C	0.7466	0.0249	rs8045689	Dixon_LCLS	2.10E-09	C	-0.508
SPNS1	3.51E-09	rs151230	2.06E-05	A	G	0.1190	0.0326	rs151230	LloydJones_WHOLEBLOOD	1.00E-10	A	-0.276115
SPNS1	3.51E-09	rs67479058	0.000161762	T	C	0.8503	-0.0258	rs67479058	Zhernakova_WHOLEBLOOD-gene-primary	3.87E-21	C	-9.43608
SPNS1	3.51E-09	rs67479058	0.000161762	T	C	0.8503	-0.0258	rs67479058	LloydJones_WHOLEBLOOD	1.70E-18	C	-0.323235
SPNS1	3.51E-09	rs67479058	0.000161762	T	C	0.8503	-0.0258	rs67479058	Zhernakova_WHOLEBLOOD-gene-primary	2.55E-13	C	-7.3162
SPNS1	3.51E-09	rs7201546	0.00291959	T	C	0.1480	0.0240	rs7201546	Westra_WHOLEBLOOD	3.16E-16	T	-8.16732
SPNS1	3.51E-09	rs7201546	0.00291959	T	C	0.1480	0.0240	rs7201546	LloydJones_WHOLEBLOOD	7.10E-12	T	-0.279907
SPNS1	3.51E-09	rs7201546	0.00291959	T	C	0.1480	0.0240	rs7201546	Zhernakova_WHOLEBLOOD-gene-primary	1.48E-11	T	-6.75016

Table E4

SPNS1	3.51E-09	rs11643913	0.0032946	A	G	0.0629	0.0519	rs11643913	LloydJones_WHOLEBLOOD	9.00E-14	A	-0.551074
SPNS1	3.51E-09	rs11569775	0.146547	C	G	0.9728	0.0206	rs11569775	LloydJones_WHOLEBLOOD	3.80E-22	G	0.71576
SPNS1	3.51E-09	rs7499778	0.193516	A	G	0.5357	-0.0073	rs7499778	Zhernakova_WHOLEBLOOD-gene-primary	8.36E-22	G	9.59551
SPNS1	3.51E-09	rs7499778	0.193516	A	G	0.5357	-0.0073	rs7499778	LloydJones_WHOLEBLOOD	1.50E-12	A	-0.195003
SPNS1	3.51E-09	rs12448482	0.197731	A	G	0.3861	0.0071	rs12448482	Westra_WHOLEBLOOD	4.65E-116	A	-22.9
SPNS1	3.51E-09	rs12448482	0.197731	A	G	0.3861	0.0071	rs12448482	LloydJones_WHOLEBLOOD	1.10E-62	A	-0.479659
SPNS1	3.51E-09	rs12448482	0.197731	A	G	0.3861	0.0071	rs12448482	Zhernakova_WHOLEBLOOD-gene-primary	5.75E-60	A	-16.3329
SPNS1	3.51E-09	rs12448482	0.197731	A	G	0.3861	0.0071	rs12448482	Zeller_MONOCYTES	2.31E-27	NA	-9
SPNS1	3.51E-09	rs12448482	0.197731	A	G	0.3861	0.0071	rs12448482	Zhernakova_WHOLEBLOOD-gene-primary	2.51E-24	A	-10.1769
SPNS1	3.51E-09	rs12448482	0.197731	A	G	0.3861	0.0071	rs12448482	Grundberg_LCLS	2.76E-10	G	0.0862493
SPNS1	3.51E-09	rs139456978	0.461272	T	C	0.0493	-0.0091	rs139456978	Zhernakova_WHOLEBLOOD-gene-primary	2.80E-19	T	-8.97634
SPNS1	3.51E-09	rs75556002	0.675961	T	C	0.0102	-0.0096	rs75556002	LloydJones_WHOLEBLOOD	1.30E-14	T	0.72673
SPNS1	3.51E-09	rs62035317	0.700585	A	G	0.0663	0.0042	rs62035317	Zhernakova_WHOLEBLOOD-gene-primary	1.12E-11	A	6.79046
SPNS1	3.51E-09	rs7498329	0.797214	A	C	0.7619	-0.0016	rs7498329	Westra_WHOLEBLOOD	1.53E-35	C	-12.443
SPNS1	3.51E-09	rs7498329	0.797214	A	C	0.7619	-0.0016	rs7498329	LloydJones_WHOLEBLOOD	4.60E-18	C	-0.286571
SPNS1	3.51E-09	rs7498329	0.797214	A	C	0.7619	-0.0016	rs7498329	Zhernakova_WHOLEBLOOD-gene-primary	2.26E-13	C	-7.33211
SPNS1	3.51E-09	rs7498329	0.797214	A	C	0.7619	-0.0016	rs7498329	Fehrmann_WHOLEBLOOD	3.10E-12	C	-6.97
SPNS1	3.51E-09	rs8056259	0.894326	T	C	0.2857	-0.0008	rs8056259	LloydJones_WHOLEBLOOD	2.50E-12	T	0.206418
SPNS1	3.51E-09	rs112918513	0.992411	A	G	0.9099	-0.0001	rs112918513	Zhernakova_WHOLEBLOOD-gene-primary	7.39E-15	G	7.77772
SULT1A1	1.96E-07	rs75539558	1.57E-06	C	G	0.6633	0.0250	rs75539558	Zhernakova_WHOLEBLOOD-exonratio-primary	3.89E-160	G	26.9642
SULT1A1	1.96E-07	rs75539558	1.57E-06	C	G	0.6633	0.0250	rs75539558	Zhernakova_WHOLEBLOOD-gene-primary	2.32E-16	G	-8.20409
SULT1A1	1.96E-07	rs75539558	1.57E-06	C	G	0.6633	0.0250	rs75539558	GTE_WHOLEBLOOD	1.14E-12	G	-0.355971
SULT1A1	1.96E-07	rs75539558	1.57E-06	C	G	0.6633	0.0250	rs75539558	GTE_SKIN	1.06E-10	G	-0.369911
SULT1A1	1.96E-07	rs231970	4.84E-06	A	G	0.8827	-0.0350	rs231970	LloydJones_WHOLEBLOOD	8.30E-25	G	-0.4597
SULT1A1	1.96E-07	rs231970	4.84E-06	A	G	0.8827	-0.0350	rs231970	Zhernakova_WHOLEBLOOD-gene-primary	3.38E-16	G	-8.15887
SULT1A1	1.96E-07	rs4788119	0.0205086	T	G	0.8112	0.0149	rs4788119	LloydJones_WHOLEBLOOD	4.60E-18	G	-0.288976
SULT1A1	1.96E-07	rs79476281	0.0875464	T	C	0.0527	-0.0218	rs150089402	Zhernakova_WHOLEBLOOD-gene-contextspecific	6.54E-13	T	-7.18876
SULT1A1	1.96E-07	rs79476281	0.0875464	T	C	0.0527	-0.0218	rs150089402	Zhernakova_WHOLEBLOOD-exon-primary	7.26E-13	T	-7.17432
SULT1A1	1.96E-07	rs79476281	0.0875464	T	C	0.0527	-0.0218	rs150089402	Zhernakova_WHOLEBLOOD-exon-primary	9.51E-12	T	-6.81362
SULT1A1	1.96E-07	rs4788115	0.122264	A	T	0.2160	0.0112	rs4788115	LloydJones_WHOLEBLOOD	4.10E-12	A	0.242542
SULT1A1	1.96E-07	rs75227850	0.159471	T	C	0.0238	-0.0191	rs75227850	LloydJones_WHOLEBLOOD	3.30E-16	T	-0.610833
SULT1A1	1.96E-07	rs11540497	0.161972	A	G	0.0493	-0.0197	rs11540497	LloydJones_WHOLEBLOOD	1.50E-10	A	-0.390543
SULT1A1	1.96E-07	rs12935321	0.533018	A	G	0.1224	0.0063	rs12445519	Zhernakova_WHOLEBLOOD-gene-primary	1.65E-29	G	11.2798
SULT1A1	1.96E-07	rs12935321	0.533018	A	G	0.1224	0.0063	rs12445519	LloydJones_WHOLEBLOOD	1.20E-10	G	0.318144
TICAM2	1.67E-06	rs17137937	2.93E-06	T	C	0.0748	0.0425	rs17137937	Hao_LUNG	0	C	-8.88
TICAM2	1.67E-06	rs17137937	2.93E-06	T	C	0.0748	0.0425	rs17137937	Westra_WHOLEBLOOD	3.67E-20	T	9.19745
TICAM2	1.67E-06	rs17137937	2.93E-06	T	C	0.0748	0.0425	rs17137937	LloydJones_WHOLEBLOOD	5.70E-10	T	0.303195
TICAM2	1.67E-06	rs256938	0.00238245	A	C	0.5646	0.0152	rs256938	Fairfax_MONOCYTES-NAIVE	2.06E-09	NA	-9
TICAM2	1.67E-06	rs2546480	0.223513	T	C	0.5527	-0.0061	rs2546480	Fairfax_MONOCYTES-IFN	3.12E-10	NA	-9
TMEM236	1.92E-06	rs7092926	6.75E-07	T	G	0.4320	0.0259	rs7092926	Zhernakova_WHOLEBLOOD-gene-primary	1.51E-26	T	-10.6634
TMEM236	1.92E-06	rs7092926	6.75E-07	T	G	0.4320	0.0259	rs7092926	Zhernakova_WHOLEBLOOD-gene-contextspecific	1.51E-26	T	-10.6634
TMEM236	1.92E-06	rs7092926	6.75E-07	T	G	0.4320	0.0259	rs7092926	Zhernakova_WHOLEBLOOD-exon-primary	4.45E-25	T	-10.3442
TMEM236	1.92E-06	rs55999004	0.00597318	A	T	0.1378	-0.0174	rs55999004	Zhernakova_WHOLEBLOOD-gene-primary	2.17E-10	A	6.34864
TMEM236	1.92E-06	rs45607131	0.0445172	T	C	0.0697	0.0178	rs45607131	Zhernakova_WHOLEBLOOD-gene-primary	2.40E-11	T	-6.67946
VPRBP	1.41E-07	rs73078636	1.41E-07	A	G	0.1480	-0.0403	rs73078636	Zhernakova_WHOLEBLOOD-exon-primary	2.14E-52	A	15.233
VPRBP	1.41E-07	rs73078636	1.41E-07	A	G	0.1480	-0.0403	rs73078636	Zhernakova_WHOLEBLOOD-gene-contextspecific	2.08E-44	A	13.9794
VPRBP	1.41E-07	rs73078636	1.41E-07	A	G	0.1480	-0.0403	rs73078636	Zhernakova_WHOLEBLOOD-gene-primary	2.08E-44	A	13.9794

* If the same eQTL appears more than once for the same gene-study-tissue combination, it indicates that the SNP was tested for association with, for example, multiple exons (RNA-seq) or multiple probes (microarray data) of the same gene.

Table E5. Summary of the directional effect of the allergy-protective allele on gene expression, across all disease-associated eQTL of a given gene.

Gene	eQTL study		N eQTL tested	N eQTL with P<0.05 in adjusted allergic disease GWAS			
				Total	Effect of allergy-protective allele on gene expression:		
	First author	Tissue			Decreased	Increased	Not available
20 genes with replication gene-based P<0.0016							
ABO	GTE	LUNG	1	0	0	0	0
ABO	GTE	SKIN	1	0	0	0	0
ABO	GTE	WHOLEBLOOD	1	1	0	1	0
ABO	Hao	LUNG	2	0	0	0	0
ABO	Yao	WHOLEBLOOD	1	1	0	1	0
ABO	Zhernakova	WHOLEBLOOD-exon-primary	1	1	0	1	0
ABO	Zhernakova	WHOLEBLOOD-exonratio-primary	1	1	0	1	0
ABO	Zhernakova	WHOLEBLOOD-gene-contextspecific	2	1	0	1	0
ABO	Zhernakova	WHOLEBLOOD-gene-primary	12	5	0	5	0
AC004893.11	Zhernakova	WHOLEBLOOD-exon-primary	1	1	1	0	0
AC004893.11	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	1	1	0	0
AC004893.11	Zhernakova	WHOLEBLOOD-gene-primary	1	1	1	0	0
APOBR	LloydJones	WHOLEBLOOD	1	1	0	1	0
APOBR	Zhernakova	WHOLEBLOOD-exon-primary	1	1	0	1	0
APOBR	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	1	0	1	0
APOBR	Zhernakova	WHOLEBLOOD-gene-primary	1	1	0	1	0
ATXN2L	Zhernakova	WHOLEBLOOD-exon-primary	1	1	1	0	0
ATXN2L	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	1	1	0	0
ATXN2L	Zhernakova	WHOLEBLOOD-gene-primary	1	1	1	0	0
FOSL2	Walsh	WHOLEBLOOD	1	1	0	1	0
FOSL2	Zhernakova	WHOLEBLOOD-exonratio-primary	1	1	0	1	0
IL27	Zhernakova	WHOLEBLOOD-exon-primary	1	1	0	1	0
IL27	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	1	0	1	0
IL27	Zhernakova	WHOLEBLOOD-gene-primary	2	2	0	2	0
IPCEF1	Battle	WHOLEBLOOD	1	1	0	0	1
IPCEF1	GTE	SKIN	1	0	0	0	0
IPCEF1	Zhernakova	WHOLEBLOOD-exon-primary	2	2	1	1	0
IPCEF1	Zhernakova	WHOLEBLOOD-exonratio-primary	1	1	1	0	0
IPCEF1	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	1	0	1	0
IPCEF1	Zhernakova	WHOLEBLOOD-gene-primary	2	2	0	2	0
LAT	Fehrmann	WHOLEBLOOD	2	2	0	1	1
LAT	Zhernakova	WHOLEBLOOD-exon-primary	1	0	0	0	0
LAT	Zhernakova	WHOLEBLOOD-gene-primary	2	1	0	1	0

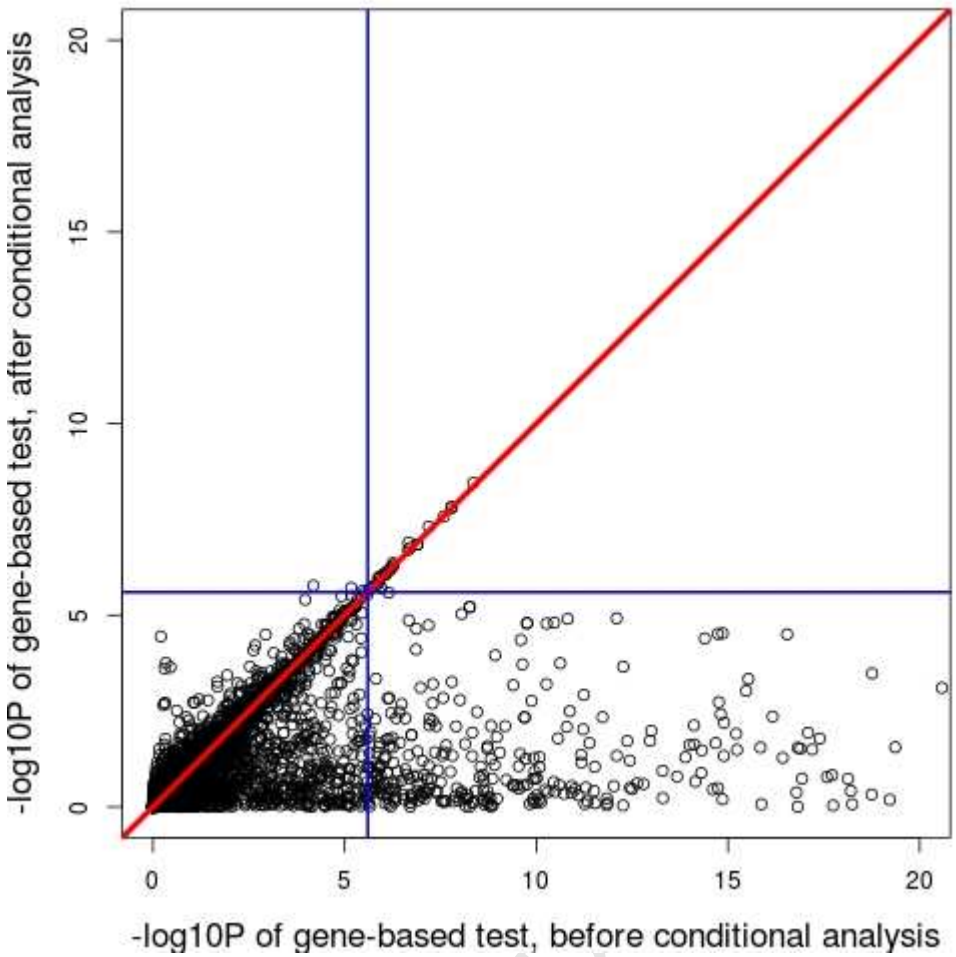
NCF4	Fairfax	MONOCYTES-LPS24	2	1	0	0	1
NCF4	Kasela	CD4TCELLS	1	0	0	0	0
NCF4	LloydJones	WHOLEBLOOD	1	1	0	1	0
NCF4	Quach	MONOCYTES-baseline	1	1	0	0	1
NCF4	Raj	CD4TCELLS	3	1	0	0	1
NCF4	Raj	MONOCYTES	1	1	0	0	1
NCF4	Westra	WHOLEBLOOD	4	2	0	2	0
NCF4	Zeller	MONOCYTES	1	1	0	0	1
NCF4	Zhernakova	WHOLEBLOOD-gene-primary	1	1	0	1	0
NSMCE1	LloydJones	WHOLEBLOOD	1	1	1	0	0
NSMCE1	Nedelec	MACROPHAGES-baseline-asQTL	1	1	0	0	1
NSMCE1	Nedelec	MACROPHAGES-listeria-asQTL	1	1	0	0	1
NSMCE1	Zeller	MONOCYTES	1	1	0	0	1
OR10J5	Westra	WHOLEBLOOD	3	3	0	3	0
PRR5L	LloydJones	WHOLEBLOOD	1	1	0	1	0
PRR5L	Zhernakova	WHOLEBLOOD-exon-primary	5	1	0	1	0
PRR5L	Zhernakova	WHOLEBLOOD-exonratio-primary	2	0	0	0	0
PRR5L	Zhernakova	WHOLEBLOOD-gene-contextspecific	2	1	0	1	0
PRR5L	Zhernakova	WHOLEBLOOD-gene-primary	3	1	0	1	0
PVALB	Battle	WHOLEBLOOD	1	1	0	0	1
PVALB	Battle	WHOLEBLOOD-ase	1	1	0	0	1
PVALB	Davenport	LEUCOCYTES	1	1	0	0	1
PVALB	Dinarzo	WHOLEBLOOD	1	1	0	1	0
PVALB	Fairfax	MONOCYTES-LPS24	1	0	0	0	0
PVALB	Fehrmann	WHOLEBLOOD	1	1	0	1	0
PVALB	GTE	WHOLEBLOOD	1	1	0	1	0
PVALB	Jansen	WHOLEBLOOD	1	1	0	1	0
PVALB	LloydJones	WHOLEBLOOD	4	2	0	2	0
PVALB	Walsh	WHOLEBLOOD	2	2	0	2	0
PVALB	Zeller	MONOCYTES	1	1	0	0	1
PVALB	Zhernakova	WHOLEBLOOD-exon-primary	2	1	0	1	0
PVALB	Zhernakova	WHOLEBLOOD-exonratio-primary	1	1	0	1	0
PVALB	Zhernakova	WHOLEBLOOD-gene-contextspecific	2	1	0	1	0
PVALB	Zhernakova	WHOLEBLOOD-gene-primary	4	2	0	2	0
PVALB	Zhernakova	WHOLEBLOOD-polyAratio-primary	1	1	0	1	0
RBM15B	Jansen	WHOLEBLOOD	1	1	0	1	0
RBM15B	Zhernakova	WHOLEBLOOD-exon-primary	1	1	0	1	0
RBM15B	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	1	0	1	0
RBM15B	Zhernakova	WHOLEBLOOD-gene-primary	1	1	0	1	0

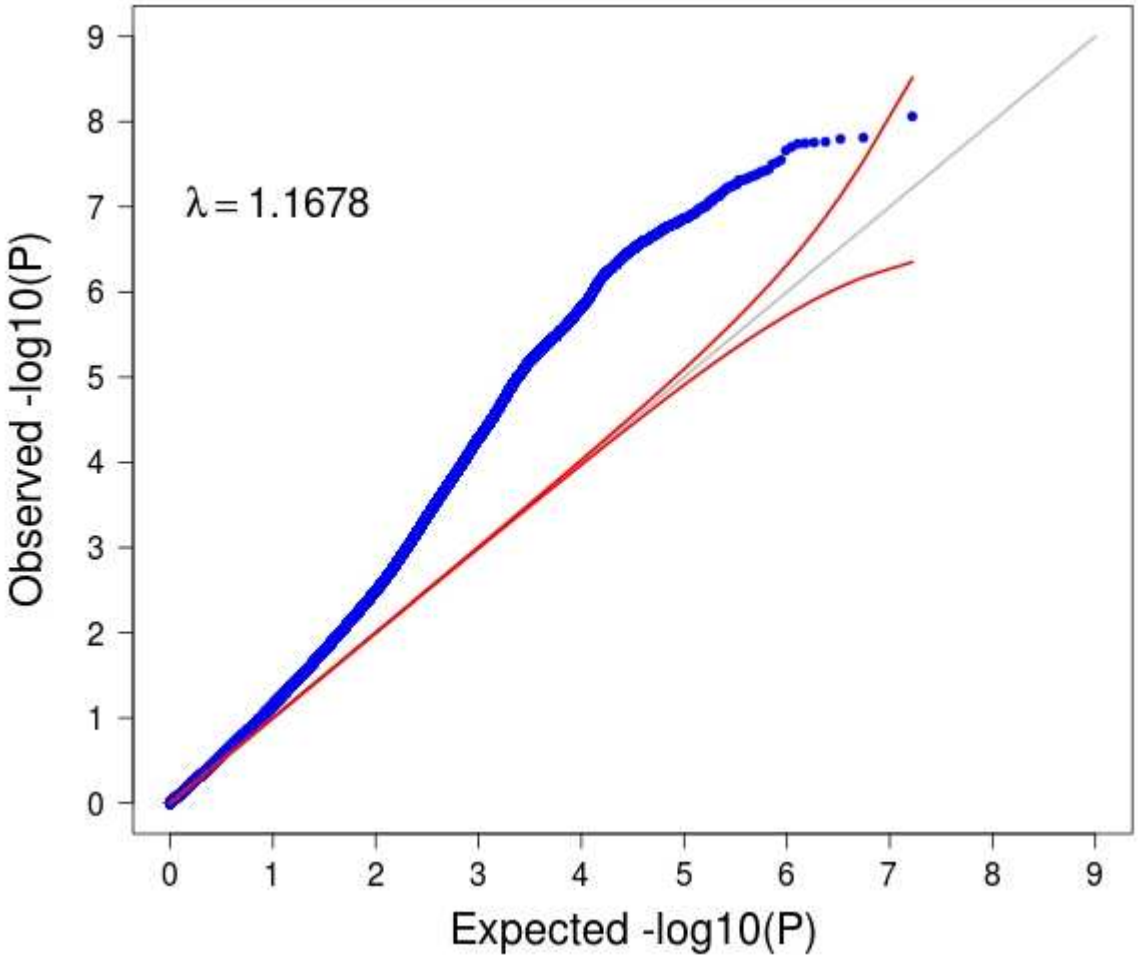
RP11-24N18.1	Zhernakova	WHOLEBLOOD-exon-primary	1	1	1	0	0
RP11-24N18.1	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	1	1	0	0
RP11-24N18.1	Zhernakova	WHOLEBLOOD-gene-primary	1	1	1	0	0
RP11-264B17.4	Zhernakova	WHOLEBLOOD-exon-primary	1	1	0	1	0
RP11-264B17.4	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	1	0	1	0
RP11-264B17.4	Zhernakova	WHOLEBLOOD-gene-primary	3	2	0	2	0
RP11-534L20.5	GTE	FIBROBLASTS	1	1	0	1	0
RP11-534L20.5	GTE	LUNG	1	1	0	1	0
RP11-534L20.5	GTE	SKIN	1	1	0	1	0
RP11-534L20.5	GTE	WHOLEBLOOD	1	1	0	1	0
RP11-534L20.5	Kasela	CD4TCELLS	4	3	0	3	0
RP11-534L20.5	Kasela	CD8TCELLS	5	3	0	3	0
RP11-534L20.5	Walsh	WHOLEBLOOD	2	1	0	1	0
RP11-534L20.5	Zhernakova	WHOLEBLOOD-gene-primary	12	5	0	5	0
SPNS1	Battle	WHOLEBLOOD	1	1	0	0	1
SPNS1	Davenport	LEUCOCYTES	1	1	0	0	1
SPNS1	Dixon	LCLS	1	1	1	0	0
SPNS1	Fairfax	MONOCYTES-LPS24	1	1	0	0	1
SPNS1	Fehrmann	WHOLEBLOOD	2	1	0	0	1
SPNS1	Geuvadis	LCLS	1	1	0	0	1
SPNS1	Grundberg	LCLS	2	1	0	1	0
SPNS1	GTE	LCLS	1	1	0	1	0
SPNS1	GTE	WHOLEBLOOD	1	1	0	1	0
SPNS1	Kim	MONOCYTES-Baseline	1	1	0	0	1
SPNS1	Kim	MONOCYTES-LPS	1	1	0	0	1
SPNS1	LloydJones	WHOLEBLOOD	12	6	0	6	0
SPNS1	Raj	CD4TCELLS	1	1	0	0	1
SPNS1	Raj	MONOCYTES	1	1	0	0	1
SPNS1	Westra	WHOLEBLOOD	4	2	0	2	0
SPNS1	Zeller	MONOCYTES	2	1	0	0	1
SPNS1	Zhernakova	WHOLEBLOOD-exon-primary	1	1	0	1	0
SPNS1	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	1	0	1	0
SPNS1	Zhernakova	WHOLEBLOOD-gene-primary	9	3	0	3	0
SULT1A1	GTE	SKIN	1	1	1	0	0
SULT1A1	GTE	WHOLEBLOOD	1	1	1	0	0
SULT1A1	LloydJones	WHOLEBLOOD	6	2	1	1	0
SULT1A1	Zhernakova	WHOLEBLOOD-exon-primary	1	0	0	0	0
SULT1A1	Zhernakova	WHOLEBLOOD-exonratio-primary	1	1	0	1	0
SULT1A1	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	0	0	0	0

SULT1A1	Zhernakova	WHOLEBLOOD-gene-primary	3	2	1	1	0
VPRBP	Zhernakova	WHOLEBLOOD-exon-primary	1	1	0	1	0
VPRBP	Zhernakova	WHOLEBLOOD-gene-contextspecific	1	1	0	1	0
VPRBP	Zhernakova	WHOLEBLOOD-gene-primary	1	1	0	1	0
3 genes with replication gene-based 0.0016<P<0.05							
CASZ1	Kim	MONOCYTES-Baseline	1	1	0	0	1
CASZ1	Kim	MONOCYTES-LPS	1	1	0	0	1
CASZ1	LloydJones	WHOLEBLOOD	1	0	0	0	0
CASZ1	Zhernakova	WHOLEBLOOD-exonratio-primary	1	0	0	0	0
FPR1	Battle	WHOLEBLOOD-ase	1	0	0	0	0
FPR1	GTE	FIBROBLASTS	1	0	0	0	0
FPR1	GTE	WHOLEBLOOD	1	1	0	1	0
FPR1	Hao	LUNG	1	1	0	1	0
FPR1	LloydJones	WHOLEBLOOD	7	3	0	3	0
FPR1	Walsh	WHOLEBLOOD	2	1	0	1	0
FPR1	Westra	WHOLEBLOOD	6	3	0	3	0
FPR1	Zeller	MONOCYTES	1	1	0	0	1
FPR1	Zhernakova	WHOLEBLOOD-exon-primary	3	1	0	1	0
FPR1	Zhernakova	WHOLEBLOOD-exonratio-primary	1	0	0	0	0
FPR1	Zhernakova	WHOLEBLOOD-gene-primary	7	6	0	6	0
TICAM2	Fairfax	MONOCYTES-IFN	1	0	0	0	0
TICAM2	Fairfax	MONOCYTES-NAIVE	1	1	0	0	1
TICAM2	Hao	LUNG	1	1	1	0	0
TICAM2	LloydJones	WHOLEBLOOD	1	1	1	0	0
TICAM2	Westra	WHOLEBLOOD	1	1	1	0	0

Table E6. Results (i.e. EUGENE P-value) from gene-based association analyses of three case-only phenotypes comparing three non-overlapping groups of adults: (1) asthma only cases (n=12,268) versus hay fever only cases (n=33,305); (2) asthma only cases (n=12,268) versus eczema only cases (n=6,276); and (3) hay fever only cases (n=33,305) versus eczema only cases (n=6,276).

Gene	Chr	BP (start)	(1) Asthma vs Hay fever	(2) Asthma vs Eczema	(3) Hay fever vs Eczema
OR10J5	1	159504793	0.0074	0.6664	0.3608
RP11-534L20.5	1	206677281	0.5574	0.2689	0.6935
FOSL2	2	28615315	0.2191	0.1755	0.1056
RBM15B	3	51428731	0.3264	0.9377	0.1260
VPRBP	3	51433298	0.3264	0.9377	0.1260
IPCEF1	6	154475631	0.2120	0.8854	0.6763
AC004893.11	7	98610788	0.9747	0.8107	0.8938
ABO	9	136125788	0.6017	0.2920	0.7406
PRR5L	11	36317838	0.0911	0.3894	0.8578
NSMCE1	16	27236312	0.6497	0.7415	0.2882
APOBR	16	28505970	0.2789	0.9077	0.4636
IL27	16	28510683	0.4716	0.9111	0.3380
SULT1A1	16	28616903	0.1836	0.9831	0.3116
ATXN2L	16	28834356	0.2188	0.9266	0.1121
RP11-24N18.1	16	28841933	0.2188	0.9266	0.1121
SPNS1	16	28985542	0.4876	0.5549	0.2796
RP11-264B17.4	16	28986294	0.5000	0.3676	0.0202
LAT	16	28996147	0.3288	0.5084	0.0163
PVALB	22	37196728	0.4415	0.2531	0.3182
NCF4	22	37257030	0.8812	0.8216	0.6894





Eleven loci with new reproducible genetic associations with allergic disease risk

ONLINE REPOSITORY

Manuel AR Ferreira, PhD [1], Judith M Vonk, PhD [2], Hansjörg Baurecht, PhD [3], Ingo Marenholz, PhD [4,5], Chao Tian, PhD [6], Joshua D Hoffman, PhD [7], Quinta Helmer, PhD [8], Annika Tillander, PhD [9], Vilhelmina Ullemar, PhD [9], Yi Lu, PhD [9], Franz Rüschendorf, PhD [4], the 23andMe Research Team[6], collaborators of the SHARE study[10], David A Hinds, PhD [6], Norbert Hübner, MD [4], Stephan Weidinger, MD[3], Patrik KE Magnusson, PhD [9], Eric Jorgenson, PhD [11], Young-Ae Lee, MD [4,5], Dorret I Boomsma, PhD [8], Robert Karlsson, PhD [9], Catarina Almqvist, MD [9,12], Gerard H Koppelman, MD [13] and Lavinia Paternoster, PhD [14]

E Tables

Tables E1 to E5 are provided in a separate Excel workbook.

Rationale for performing the conditional analysis prior to the gene-based analysis

The goal of our gene-based analysis was to identify genes with a significant association with allergic disease risk that was not driven by the 136 sentinel allergy risk SNPs reported by Ferreira et al. [1]. In practice, applying the gene-based test to the original results of Ferreira et al. [1] (i.e. before conditional analysis) would not be the best way to achieve that goal. Had we done so, there would be hundreds (exactly 362) of genes with a significant association at the genome-wide significance threshold of $P<2.5x10^{-6}$ (Figure E2). Most of the 362 genes are located in close proximity to (<1 Mb) the 136 sentinel allergy risk SNPs that we reported by Ferreira et al. [1], and so the observed gene-based associations are very likely to be driven by those known allergy risk SNPs. As indicated above, the goal of our analysis was not to identify such genes, but instead those with a gene-based association driven by potentially new allergy risk SNPs.

To do so, we could have simply filtered out genes that were located "close" to the 136 sentinel risk SNPs reported by Ferreira et al. [1]. But that is not an ideal approach, because LD patterns are not uniform throughout the genome, with substantial LD between variants found at distances of >1 Mb in some regions [2]. So if we had defined "close" as <1 Mb from a sentinel variant, then we could potentially fail to filter out genes located at a greater distance but still with an association driven by a known allergy risk SNP. Conversely, had we used a distance threshold that would ensure linkage equilibrium across the genome between the Ferreira et al. [1] sentinel risk SNPs and eQTL (e.g. >10 Mb), we would potentially encounter the opposite problem, and filter out genes with an association not driven by a known allergy risk SNP.

Another approach could have been to remove results for the 136 sentinel risk SNPs (and for any other variants in LD with those) from the Ferreira et al. [1] GWAS. Then apply the gene-based test after those

exclusions. In our view, this is more appropriate than the "distance-based" approach described above, but still not ideal, because it requires defining a threshold to identify variants in LD with the sentinel risk SNPs. Should this be an r^2 of 0.10, 0.05, or 0.01? The threshold selected could in practice turn out to be too liberal or too conservative.

Instead of the distance- or LD-based filtering approaches described above, we argue that a more efficient (and still simple and robust) approach to identify genes with an association not driven by a sentinel risk SNP identified in Ferreira et al. [1], is to first adjust the single-SNP results of that GWAS for the effects of the sentinel risk SNPs identified in that GWAS. Then apply the gene-based test to the adjusted GWAS results. Approximate conditional association analysis can be (and was in our study) performed with the GCTA tool [2], requiring only summary statistics from the GWAS and genotype data from individuals with an ancestry that is representative of the studies that were included in the GWAS (we used 5,000 Europeans from the UK Biobank).

In practice, the *P*-value for any variant located >10 Mb from a sentinel risk SNP is the same before and after the conditional analysis performed with GCTA, because both variants are considered to be in linkage equilibrium. For variants located <10 Mb from a sentinel variant, the *P*-value after the conditional analysis in GCTA will be attenuated if in (and in proportion to the) LD with the nearby sentinel risk SNP(s). At one extreme (when r^2=0), the *P*-value will be (almost if not) identical before and after conditional analysis. On the other hand, if the r^2 between the test variant and the sentinel variant is 1, then the *P*-value after conditioning on the sentinel variant will be 1 (or very close to).

When we applied the gene-based test to the adjusted GWAS results, we identified 30 genes with a $P<2.5\text{x}10^{-6}$. For most of these genes, the gene-based *P*-value obtained with the original GWAS results

(i.e. prior to the conditional analysis) was virtually unchanged (see Figure E2), confirming that the conditional analysis had not introduced biases that affected our results.

Regarding the observed inflation of test statistics after the conditional analysis (lambda of 1.17), it is important to note that lambda was greater (1.19) in our original GWAS[1] (i.e. prior to the conditional analysis). Therefore, the conditional analysis did not cause a systematic inflation of test statistics. We showed in Ferreira et al. [1] that most of this inflation is likely to reflect the polygenic nature of allergic disease risk and the large sample size used, which results in many thousands of SNPs being truly associated with disease risk (which increases lambda). This is likely to be the case because the inflation due to technical biases was only 1.04 (intercept from LD score regression analysis [2]). The Ferreira et al. GWAS results were adjusted by this inflation prior to the gene-based analysis reported in our current study.

Collaborators of the SHARE study

Jorge Esparza-Gordillo[1,2,3], Oliver Hummel[1], Sarah Grosche[1,2], John S Witte[4], Jouke-Jan Hottenga[5], Gonneke Willemsen[5], Elke Rodríguez[9], Melanie Hotze[9], Andre Franke[10], Melanie C Matheson[11], Shyamali C Dharmage[11], Andreas Arnold[12], Georg Homuth[13], Carsten O Schmidt[14], Philip J Thompson[15], Nicholas G Martin[16], David L Duffy[16], Natalija Novak[17], Holger Schulz[18,19], Stefan Karrasch[18,20], Christian Gieger[21], Konstantin Strauch[22,23], Ronald B Melles[24] and David A Hinds[25]

1 Max Delbrück Center (MDC) for Molecular Medicine, Berlin, Germany

2 Clinic for Pediatric Allergy, Experimental and Clinical Research Center of Charité Universitätsmedizin Berlin and Max Delbrück Center, Berlin, Germany

3 Current address: GlaxoSmithKline, Stevenage, UK

4 Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA

5 Department Biological Psychology, Netherlands Twin Register , Vrije University, Amsterdam, The Netherlands

9 Department of Dermatology, Allergology and Venereology, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany

10 Institute of Clinical Molecular Biology, Christian Albrechts University of Kiel, Kiel, Germany

11 Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia

12 Clinic and Polyclinic of Dermatology, University Medicine Greifswald, Greifswald, Germany

13 Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany

14 Institute for Community Medicine, Study of Health in Pomerania/KEF, University Medicine

Greifswald, Greifswald, Germany

15 Institute for Respiratory Health, Harry Perkins Institute of Medical Research, University of Western Australia, Nedlands, Australia

16 Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Australia

17 Department of Dermatology and Allergology, University-Hospital Bonn, Bonn, Germany

18 Institute of Epidemiology I, Helmholtz Zentrum Munchen - German Research Center for Environmental Health , Neuherberg, Germany

19 Comprehensive Pneumology Center Munich (CPC-M), Member of the German Center for Lung Research, Munich, Germany

20 Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Ludwig-Maximilians-Universität, Munich, Germany

21 Research Unit of Molecular Epidemiology and Institute of Epidemiology II,, Helmholtz Zentrum Munchen - German Research Center for Environmental Health , Neuherberg, Germany

22 Institute of Genetic Epidemiology, Helmholtz Zentrum Munchen - German Research Center for Environmental Health , Neuherberg, Germany

23 Chair of Genetic Epidemiology, IBE, Faculty of Medicine, LMU Munich, Germany

24 Division of Research, Kaiser Permanente Northern California, Oakland, California, USA

25 23andMe, Inc., Mountain View, California, USA

Acknowledgments

23andMe: We would like to thank the research participants and employees of 23andMe for making this work possible. We particularly thank the following members of the 23andMe Research Team: Michelle Agee, Babak Alipanahi, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah L. Elson, Pierre Fontanillas, Nicholas A. Furlotte, Bethann S. Hromatka, Karen E. Huber, Aaron Kleinman, Nadia K. Litterman, Jennifer C. McCreight, Matthew H. McIntyre, Joanna L. Mountain, Elizabeth S. Noblin, Carrie A.M. Northover, Steven J. Pitts, J. Fah Sathirapongsasuti, Olga V. Sazonova, Janie F. Shelton, Suyash Shringarpure, Joyce Y. Tung, Vladimir Vacic, and Catherine H. Wilson.

LifeLines: The LifeLines Biobank initiative has been made possible by funds from FES (Fonds Economische Structuurversterking), SNN (Samenwerkingsverband Noord Nederland) and REP (Ruimtelijk Economisch Programma). The authors wish to acknowledge the services of the LifeLines Cohort Study, the contributing research centres delivering data to LifeLines, and all the study participants.

SALTY/TWINGENE/CATSS: We acknowledge funding from the Swedish Research Council. Financial support was provided by the Swedish Research Council through the Swedish Initiative for research on Microdata in the Social and Medical Sciences (SIMSAM) framework grant number 340-2013-5867 and grants from the Swedish Heart-Lung Foundation. Computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX.

ALSPAC: The UK Medical Research Council and the Wellcome Trust (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. ALSPAC GWAS data was generated by

Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe.

AAGC: The AAGC was funded by a grant from the NHMRC (project ID 613627).

GENEVA: The project received infrastructure support through the DFG Clusters of Excellence "Inflammation at Interfaces" (grants EXC306 and EXC306/2), and was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (sysINFLAME, grant # 01ZX1306A), and the PopGen 2.0 network (01EY1103). The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

NTR: This study was supported by multiple grants from the Netherlands Organization for Scientific Research (NWO: 016-115-035, 463-06-001, 451- 04-034); ZonMW (31160008, 911-09-032); and NWO 480-15-001/674: Netherlands Twin Registry Repository: researching the interplay between genome and environment; The Amsterdam Public Health Institute (APH) and Neuroscience Campus Amsterdam (NCA); Biomolecular Resources Research Infrastructure (BBMRI–NL, 184.021.007), European Research Council (ERC-230374); Genotyping was made possible by grants from NWO/SPI 56-464-14192, Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health, Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06), the Avera Institute, Sioux Falls (USA) and the National Institutes of Health (NIH R01 HD042157-01A1,

MH081802, Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995).

GENUFAD-SHIP-1.We thank all individuals and families for their participation in this study. We thank all physicians and nurses involved in patient recruitment for their valuable contribution to the study. We are grateful to the laboratory technicians C. Flachmeier and T. Thuss for their work. The study was funded by the German Ministry of Education and Research (BMBF) through the Clinical Research Group for Allergy at Charité Berlin, the National Genome Research Network (NGFN). The SHIP authors are grateful to Mario Stanke for the opportunity to use his server cluster for SNP imputation. We thank all staff members and participants of the SHIP studies, as well as all of the genotyping staff for generating the SHIP SNP data set. SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network ‘Greifswald Approach to Individualized Medicine (GANI_MED)’ funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data were supported by the Federal Ministry of Education and Research (grant 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg–West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG and the Caché Campus program of the InterSystems GmbH.

GENUFAD-SHIP-2. We thank all individuals and families for their participation in this study. We thank all physicians and nurses involved in patient recruitment for their valuable contribution to the study. We are grateful to the laboratory technicians C. Flachmeier and T. Thuss for their work. The study was funded by the German Ministry of Education and Research (BMBF) through the Clinical Research

Group for Allergy at Charité Berlin, the National Genome Research Network (NGFN). The SHIP authors are grateful to Mario Stanke for the opportunity to use his server cluster for SNP imputation. We thank all staff members and participants of the SHIP studies, as well as all of the genotyping staff for generating the SHIP SNP data set. SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network ‘Greifswald Approach to Individualized Medicine (GANI_MED)’ funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data were supported by the Federal Ministry of Education and Research (grant 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg–West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG and the Caché Campus program of the InterSystems GmbH.

References

1. Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. Nat Genet 2017; 49:1752-7.
2. Yang J, Ferreira T, Morris AP, Medland SE, Madden PA, Heath AC, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 2012; 44:369-75, S1-3.

Figure E1. Distribution of the observed and expected single-SNP association P-values obtained after adjusting the results from Ferreira et al. [1] (which included a correction for an inflation factor of 1.04; see Methods for details) for the effects of the 136 independent associations identified in that study. The genomic inflation factor (λ, estimated as the median chi-square divided by 0.4549) of the single-SNP results after the conditional analysis is also shown (λ = 1.17); before the conditional analysis, λ was 1.19. The intercept of LD-score regression was 1.00, both before and after the conditional analysis.

Figure E2. Comparison of results from the EUGENE gene-based test obtained before (x-axis) and after (y-axis) adjusting the Ferreira et al.[1] GWAS results for the effects of the sentinel risk SNPs identified in that GWAS. The vertical and horizontal blue lines indicate the P-value threshold used to identify significant associations after accounting for multiple testing (P=2.5x10^{-6}).